AD_____

Award Number: W81XWH-05-1-0277

TITLE: Regulated Ubiquitinylation-Dependent Corepressor and Coactivator Complex Exchange as a Potential Target in Estrogen-Dependent Breast Cancer

PRINCIPAL INVESTIGATOR: Christine Shiang-Ling Cheng

CONTRACTING ORGANIZATION: University of California
San Diego, CA 92093-0934

REPORT DATE: September 2006

TYPE OF REPORT: Annual Summary

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
01-09-2006	Annual Summary	1 Sep 2005 – 31 Aug 2006

4. TITLE AND SUBTITLE

Regulated Ubiquitinylation-Dependent Corepressor and Coactivator Complex Exchange as a Potential Target in Estrogen-Dependent Breast Cancer

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-05-1-0277

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Christine Shiang-Ling Cheng

E-Mail: c3cheng@ucsd.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of California
San Diego, CA 92093-0934

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES - original document contains color figures.

14. ABSTRACT

The initial focus of my first year of training has been on getting training in basic molecular biology experimental techniques and bioinformatics analysis of microarray and genome-wide chromatin immunoprecipitation data. Successful initial research results have been accepted for publication. Specific Aim 1) Mutant constructs that block specific phosphorylation sites in TBL1 and TBLR1 has been constructed for microinjection and trasnfection experiments totest functionality of each phosphorylation sites. Phosphopeptide specific antibodies for TBL1 and TBLR1 have been developed and initial characterization of the antibody has been performed. Specific Aim 2) Performed data analysis of genome-wide chromatin immunoprecipitation experiments with ERα, RNA polII and histone modification markers and correlate the binding data with expression profiles upon estrogen stimulation.

15. SUBJECT TERMS
No subject terms provided

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			USAMRMC
U	U	U	UU	29	**19b. TELEPHONE NUMBER** *(include area code)*

Table of Contents

Introduction

This is my first year of training upon the recipient of the Predoctoral Traineeship Award from the Army Breast Cancer Research Program. The focus has been mainly on getting training in basic molecular biology experimental techniques and bioinformatics analysis of microarray and genome-wide chromatin immunoprecipitation data. Successful initial research results have been accepted for publication. I am now able to apply my skills to a new interdisciplinary research project where the role of NFκB in endocrine-resistant breast cancer will be studied using microarray dataset and computational models.

Body

The initial training has been focused on receive training on experimental techniques in molecular biology and bioinformatics skills to analyze genome scale datasets. I have received training and have demonstrated the ability to perform molecular biology experimental techniques, such as cloning, western blot, luciferase reporter assay, immunohistochemistry, RT-PCR, chromatin immunoprecipication and chromosome conformational capture assay. I have also learned to conduct bioinformatics analysis of expression profiles from microarray dataset, transcription factor binding dataset from genome-wide chromatin immunoprecipication experiments, clustering analysis, gene ontology analysis and motif search.

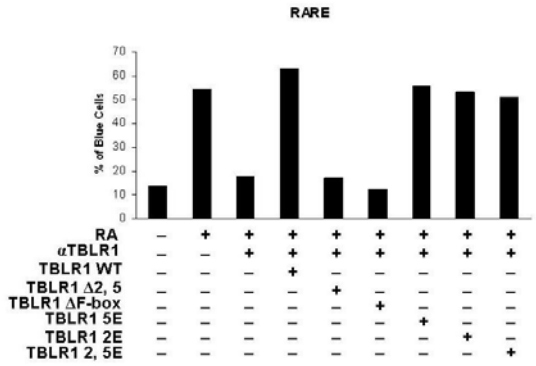

Figure 1. Specific phosphorylation sites on TBLR1 are required for ligand-dependent transcriptional activation. Inhibition of ligand dependent transcriptional activation by antibodies against TBLR1 was rescued by co-injection of expression vector for wt TBLR1, while inhibition could not be rescued when the phosphorylation sites on TBLR1 were mutated to block phosphorylation.

The first specific aim of the project has been on understanding the potential roles for specific phosphorylation sites in TBL1 and TBLR1 in estrogen-dependent breast cancer. Additional phosphorylation mutants (serine to alanine and serine to glutamic acid mutations) of TBL1 and TBLR1 have been constructed. Microinjection experiments examining the effects of each of these mutants upon estrogen and tamoxifen treatment have been

performed and have further demonstrated the requirement for these phosphorylation sites in ligand dependent transcriptional activation event (Figure 1). Interestingly, one of the phosphorylation site on TBLR1 might have a role in transcriptional repression as the basal transcriptional activity was further attenuated when the expression vector of TBLR1 phosphorylation mutant that mimic phosphorylation (serine to glutamic acid mutation) were microinjected (Figure 2). Initial characterization of the phosphopeptide specific antibodies of TBL1 and TBLR1 has been performed. Initially some results from transfection experiments and microinjection experiments have not been completely in agreement with each other. After careful examination, we have identified the difference was due to a specific bottle of cultural medium (fetal calf serum) that might have been too old that gave inconsistent results and was used only in some of the transfection experiments. With new batch of serum, the results are now consistent between the two assays and are consistent with our initial results.

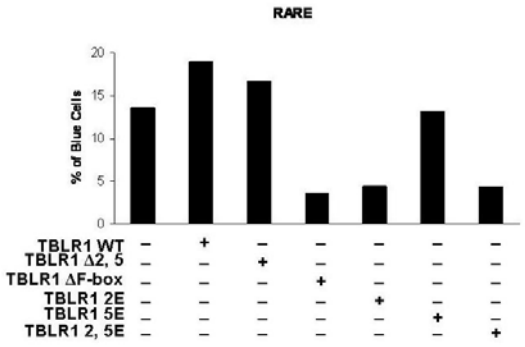

Figure 2. One of the specific phosphorylation sites on TBLR1 (2E) may be involved in ligand-dependent transcriptional repression. Basal transcriptional activity was further attenuated when expression vector of TBLR1 2E mutant, a serine to glutamic acid mutation that potentially mimic the phosphorylation, was microinjected.

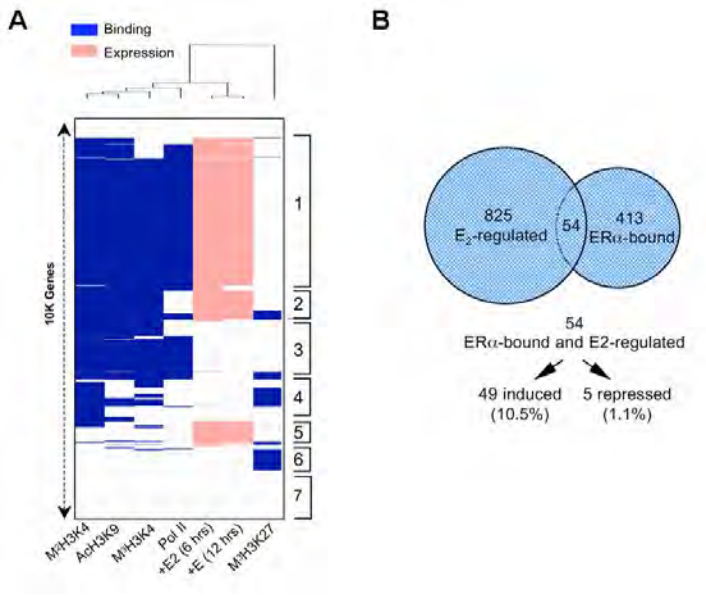

Figure 3. (A) Correlation of gene expression with promoter occupancy by Pol II and histone modification markers. Gene expression profiling in E_2-induced MCF-7 cells was carried out on Illumina gene expression arrays. ~10, 000 genes common to both promoter and expression profiling arrays and reliable scores in all measurements were used to construct the binary map by unsupervised hierarchical clustering analysis. Gene clusters are grouped based on the overall pattern. (B) Venn diagram showing the overlap between ER -bound promoters and E_2-induced genes.

The second specific aim of the proposed project was to perform genome-wide chromatin immunoprecipitation experiments with phosphopeptide specific antibodies of TBL1 and TBLR1. A prerequisite step is to identify ERα binding targets by performing ERα genome-wide chromatin immunoprecipitation and also to understand which of these ERα binding targets are regulated in response to estrodial. Once the ERα transcriptional regulatory program has been defined, we can then utilize the same platform, analysis tools and the knowledge about ERα further with TBL1 and TBLR1 to define the binding and functional involvement of each specific phosphorylation sites. Thus my initial effort has been focused on analyzing genome-wide chromatin immunoprecipitation experiments with ERα, RNA polymerase II and different histone markers and how these transcription factor binding data correlate with the expression profile upon estrogen induction. More detailed description of the data analysis and experiments and discussion of the results are presented in the two journal articles that have been published (Appendices A and B).

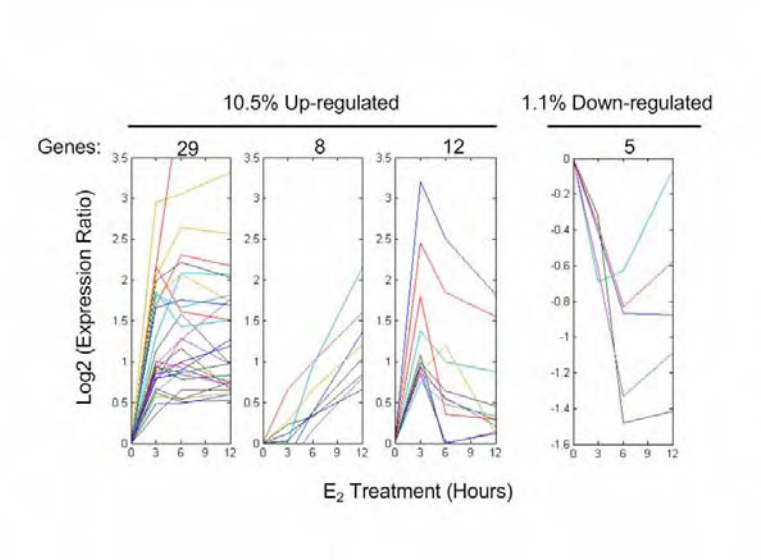

Figure 4. Gene expression profiling in response to E_2-treatment. ERα-bound and E_2-regulated genes are grouped into four distinct classes. Among up-regulated genes, 29 were rapidly induced and the level remained relatively constant afterward, 8 were induced in a time-dependent manner; and 12 were induced followed by a rapid decay. E_2-induced genes represent 10.5% of total ERα-bound genes in the promoter-proximal region. Only 5 ERα-bound genes were down-regulated by E_2, which represent 1.1% of total ERα-bound genes in the promoter-proximal region.

Correlation of gene expression with promoter occupany by PolII and histone modification markers was visualized by constructing the binary map using hierarchical clustering analysis (Figure 3). ERα bound promoters were identified at $p < 0.0001$ in E_2-induced MCF7 cells by a novel, highly sensitive genome-wide chromatin immunoprecipitation method, ChIP-DSL. We evaluated the time-course of regulated gene expression by RNA profiling and identified 879 genes that responded to E_2-

induction in MCF7 cells, which generally agrees with other published gene expression profiling studies. Strikingly, only 54 out of these 879 E_2-affected genes were bound by ERα in the promoter proximal region (Figure 3), indicating that the majority of E_2-induced genes might be indirectly affected or regulated by ER responsive elements (EREs) located away from the promoter proximal region. Among these 879 E_2-regulated genes, 562 were up-regulated and 317 were down-regulated. ERα-bound and E_2-regulated genes are grouped into four distinct classes. Among up-regulated genes, 29 were rapidly induced and the level remained relatively constant afterwards; 8 were induced in a time-dependent manner; and 12 were induced followed by a rapid decay. E_2-induced genes represent 10.5% of total ERα-bound genes in the promoter-proximal region. Only 5 ERα-bound genes were down-regulated by E_2, which represent 1.1% of total ERα-bound genes in the promoter-proximal region (Figure 4). We also have discovered that a large portion of the ERα binding genes were not regulated upon estrogen treatment.

Further analysis based on microarray data from the literature has helped us targeted to a set of genes that have ERα binding and were not regulated by E_2 in MCF7 cells but were actually induced in U2OS cells. Chromatin immunoprecipitation experiments were performed with ERα antibody in U2OS cells and found ERα were binding to the promoter of these genes. Quantitative RT-PCR were also performed to further validate that these gene are induced by

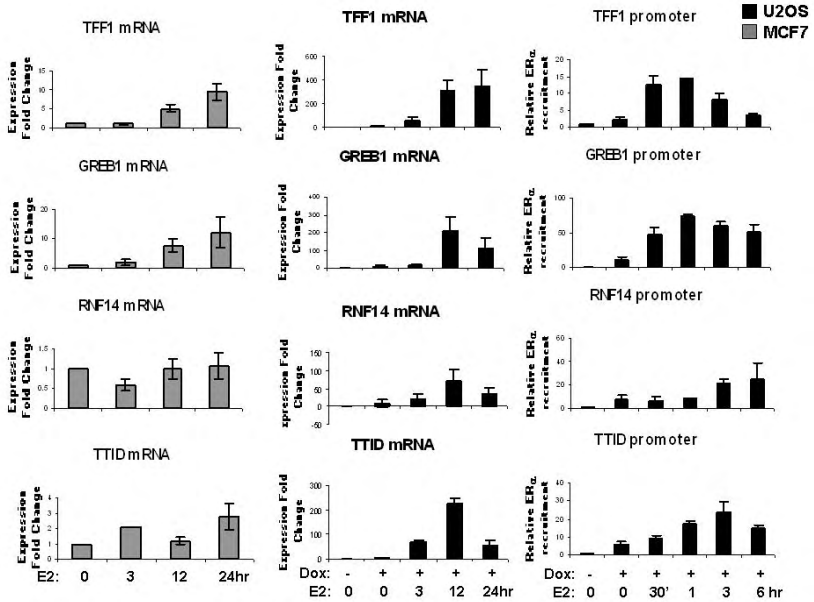

Figure 5. Quantitative RT-PCR showed induced levels of RNF14 and TTID mRNA in response to E_2 in U2OS cells but not in MCF-7 cells. ERα recruitment to the promoter of RNF14 and TTID in U2OS cells were demonstrated by chromatin immunoprecipitation assay. TFF1 and GREB1 represent genes that are induced by E_2 in both MCF-7 cells and U2OS cells. ERα also recruit to the promoter of TFF1 and CREB1 in U2OS cells.

estrogen in the mRNA level (Figure 5). These results demonstrated that ERα occupancy on the promoter does not directly translates to induction of the gene upon estrogen stimulation. ERα target genes may be induced only under certain cellular condition, but ERα binding on the promoter potentiate the gene to respond to specific signaling stimulus. To further investigate the biological relevance of ERα-binding and estrogen-regulated gene expression, we asked how different ERα target genes might be differentially regulated in breast cancer tissues using a comprehensive set of gene expression profiling data from 251 breast cancer patients. Focusing on ERα-bound and E_2 up-regulated genes, we found a direct correlation between gene expression and tumor progression by unsupervised hierarchical clustering. Patients were clustered into three groups. About half of the genes were strongly suppressed in group2, which displayed an ER negative status and advanced tumor grade (Figure 6). Significantly, this patient group exhibited a much reduced survival rate compared to the two other groups (Figure 6).

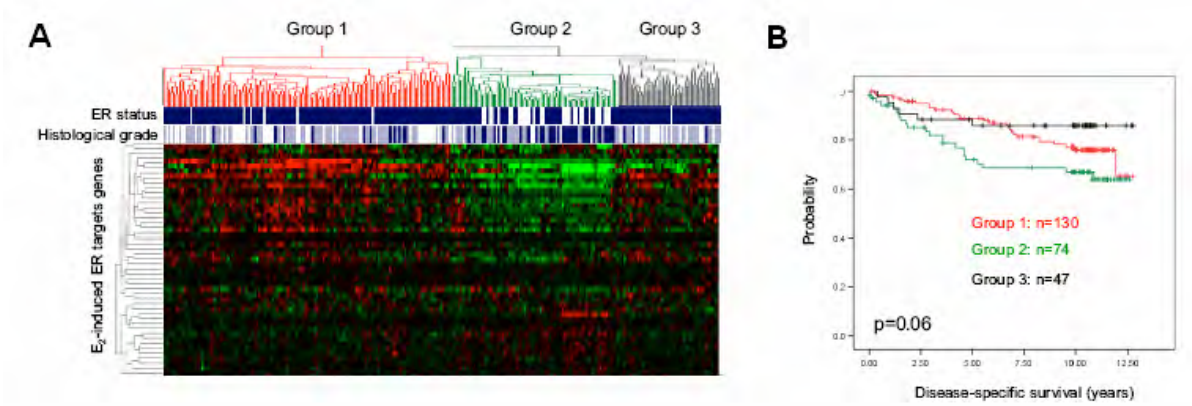

Figure 6 (A) Segregation of ER expression and breast tumor grade (both indicated on top by blue bars) based on ERα-bound, E_2-induced genes in MCF-7 cells. (B) Kaplan-Meier plots of patient survival in different groups segregated based on ER binding and E_2-induction. Statistical significance was determined by χ^2 test.

Key Research Accomplishments

1. **Comparison of genome-wide chromatin immunoprecipitation experiment with ERα and RNA profiling upon estradiol stimulation revealed that a large proportion of ERα-binding genes are not regulated by estradiol in MCF-7 cells.**

2. **Chromatin immunoprecipitation experiment with ERα and expression profiling with RT-PCR of selected genes upon estradiol stimulation in U2OS cells revealed some ERα-binding genes that are not regulated by estradiol in MCF-7 cells were in fact bound by ERα but are induced by estradiol in U2OS cells implicating a possibility that the expression of ERα target genes being regulated by estradiol only when a certain physiological cellular/signaling condition is met.**

Reportable Outcomes

1. **Garcia-Bassets I, Kwon YS, Telese F, Prefontaine GG, Hutt KR, Cheng CS, Ju BG, Ohgi KA, Wang J, Escoubet-Lozach L, Rose DW, Glass CK, Fu XD, Rosenfeld MG. Histone methylation-dependent mechanisms impose ligand dependency for gene activation by nuclear receptors. Cell. 2007 Feb 9;128(3):505-18. (Appendices A)**

2. **Kwon YS, Garcia-Bassets I, Hutt KR, Cheng CS, Jin M, Liu D, Benner C, Wang D, Ye Z, Bibikova M, Fan JB, Duan L, Glass CK, Rosenfeld MG, Fu XD. Sensitive ChIP-DSL technology reveals an extensive estrogen receptor alpha-binding program on human gene promoters. Proc Natl Acad Sci U S A. 2007 Mar 20;104(12):4852-7. (Appendices B)**

References

Miller, L. D., Smeds, J., George, J., Vega, V. B., Vergara, L., Ploner, A., Pawitan, Y., Hall, P., Klaar, S., Liu, E. T., and Bergh, J. (2005). An expression signature for p53 status in human breast cancer predicts mutation status, transcriptional effects, and patient survival. Proc Natl Acad Sci U S A 102, 13550-13555.

Cell

Histone Methylation-Dependent Mechanisms Impose Ligand Dependency for Gene Activation by Nuclear Receptors

Ivan Garcia-Bassets,[1,4] Young-Soo Kwon,[2,4] Francesca Telese,[1] Gratien G. Prefontaine,[1] Kasey R. Hutt,[1] Christine S. Cheng,[1] Bong-Gun Ju,[1] Kenneth A. Ohgi,[1] Jianxun Wang,[1] Laure Escoubet-Lozach,[2] David W. Rose,[3] Christopher K. Glass,[2,*] Xiang-Dong Fu,[2,*] and Michael G. Rosenfeld[1,*]
[1] Howard Hughes Medical Institute, Department of Medicine
[2] Department of Cellular and Molecular Medicine
[3] Department of Medicine, Division of Endocrinology and Metabolism
University of California, San Diego, School of Medicine, 9500 Gilman Drive, La Jolla, CA 92093, USA
[4] These authors contributed equally to this work.
*Correspondence: ckg@ucsd.edu (C.K.G.), xdfu@ucsd.edu (X.-D.F.), mrosenfeld@ucsd.edu (M.G.R.)
DOI 10.1016/j.cell.2006.12.038

SUMMARY

Nuclear receptors undergo ligand-dependent conformational changes that are required for corepressor-coactivator exchange, but whether there is an actual requirement for specific epigenetic landmarks to impose ligand dependency for gene activation remains unknown. Here we report an unexpected and general strategy that is based on the requirement for specific cohorts of inhibitory histone methyltransferases (HMTs) to impose gene-specific gatekeeper functions that prevent unliganded nuclear receptors and other classes of regulated transcription factors from binding to their target gene promoters and causing constitutive gene activation in the absence of stimulating signals. This strategy, based at least in part on an HMT-dependent inhibitory histone code, imposes a requirement for specific histone demethylases, including LSD1, to permit ligand- and signal-dependent activation of regulated gene expression. These events link an inhibitory methylation component of the histone code to a broadly used strategy that circumvents pathological constitutive gene induction by physiologically regulated transcription factors.

INTRODUCTION

Ligand-dependent activation of large transcriptional gene programs that are simultaneously regulated by nuclear receptors provides critical strategies for development and homeostasis of all metazoans, and its misregulation is associated with many types of disease. Regulated transcription by nuclear receptors is mediated by ligands binding to the C-terminal domain, thus causing conformational changes; these include a change in the position of the so-called AF2 helix, which favors association with specific coactivator complexes and functional conversion of the receptor to an activator (reviewed in Rosenfeld et al., 2006). Thus, when unliganded, nuclear receptors, such as the thyroid hormone (T_3) and the retinoid acid (RA) receptors, act as repressors primarily by recruiting specific corepressor complexes via the "CoRNR" domain (Horlein et al., 1995; Chen and Evans, 1995; Heinzel et al., 1997; Privalsky, 2004), but, when liganded, they are functionally converted to activators by recruiting coactivator complexes. In addition, for many nuclear receptors, such as estrogen receptor α (ERα) and androgen receptor (AR), other signaling pathways can cause similar recruitment of coactivators and the consequent functional conversion to transcriptional activators even in the absence of ligand (Culig et al., 1994; Nazareth and Weigel, 1996; Zwijsen et al., 1997; Rogatsky et al., 1999; Ueda et al., 2002; Ogawa et al., 2004; Kim et al., 2005). Therefore, it is of particular interest to further explore the linkage between the recruitment of nuclear receptors and the coregulatory complexes that underlie ligand-dependent and -independent activation of transcriptional programs.

Most coregulatory complexes exhibit a diversity of enzymatic activities that can be divided into two generic classes: enzymes capable of remodeling the structure of the nucleosome in an ATP-dependent manner and enzymes capable of covalently modifying histone tails; this latter group includes acetylating and deacetylating activities (HATs and HDACs); methylating and demethylating activities (HMTs and HDMs); kinases and phosphatases; poly(ADP)ribosylases; and ubiquitin and SUMO ligases (reviewed in Narlikar et al., 2002; Rosenfeld et al., 2006). The histone code model (Strahl and Allis, 2000; Jenuwein and Allis, 2001) suggests that serial posttranslational histone modifications, such as acetylation, methylation,

phosphorylation, and sumoylation, correlate with the specific activated or repressed status of the promoter (reviewed in Fischle et al., 2003; Peterson and Laniel, 2004; Margueron et al., 2005). Presumably, the deposition and removal of these marks correspond to a large variety of coregulatory complexes that act in a sequential and combinatorial fashion to ultimately determine spatial and temporal control of gene expression (reviewed in McKenna and O'Malley, 2002; Rosenfeld et al., 2006). One of these marks, the histone lysine methylation, was initially considered as a permanent posttranslational modification that exerts long-term epigenetic memory (reviewed in Kouzarides, 2002; Lachner and Jenuwein, 2002). However, recent data demonstrating the existence of lysine demethylase activities have dramatically challenged this model (Shi et al., 2004; Metzger et al., 2005; Tsukada et al., 2006; Yamane et al., 2006; Whetstine et al., 2006).

Histone lysine methylation has been extensively linked to both gene activation and gene repression events in euchromatic and heterochromatic regions (reviewed in Lachner and Jenuwein, 2002; Martin and Zhang, 2005). A large number of SET-domain-containing enzymes, including RIZ1, ESET, Eu-HMTase1, G9a, Suv39h1/h2, MLL1, and others, have been shown to transfer methyl groups to histones and to transcription factors; in particular, this has been shown at multiple lysine residues in histones, including H3-K4, K9, K27, K36, K79, H4-K20, and H1-K26, all of which have been reported in most cases to be mono-, di-, and trimethylated (reviewed in Martin and Zhang, 2005). It has been proposed that methyl groups may act as binding sites for a wide range of chromatin proteins, including the repressive heterochromatin protein 1 (HP1), which has been reported to bind methyl groups on histone H3 at lysine 9 (H3-K9; Nielsen et al., 2001), as well as the transcriptional activator WDR5 and the ATP-dependent chromatin-remodeling protein CDH1, which have been reported to bind methyl groups on H3 at lysine 4 (H3-K4; Wysocka et al., 2005; Dou et al., 2005; Flanagan et al., 2005).

Recently, a CoREST corepressor complex component (Tong et al., 1998; Andres et al., 1999; Ballas et al., 2001; Humphrey et al., 2001; You et al., 2001; Hakimi et al., 2002, 2003; Lunyak et al., 2002; Shi et al., 2003) known as LSD1/BHC110/KIAA0601/p110b was identified as the first histone lysine demethylase, which is involved in mediating neuron-restrictive silencing factor (NRSF)/REST-dependent repression of neuronal genes in non-neuronal cells (Shi et al., 2004, 2005). In vitro, LSD1 specifically demethylates mono- and dimethylated H3-K4, which, in vivo, has been suggested to mediate gene repression by maintaining an unmethylated H3-K4 status on a specific set of regulated promoters (Shi et al., 2004, 2005). Interestingly, the demethylase activity of LSD1 has been shown to be modulated by other proteins, including CoREST and BHC80 (Shi et al., 2005; Lee et al., 2005), and by other histone marks displayed around the substrate (Forneris et al., 2005, 2006). Furthermore, LSD1 has been implicated in H3-K9 demethylation that is associated with *PSA* gene activation in an AR-dependent manner (Metzger et al., 2005) and has been described as a component of the MLL1 coactivator complex (Nakamura et al., 2002). More recently, additional demethylase enzymes have been described and include JHDM1, which specifically demethylates dimethyl histone H3 at lysine 36 (H3-K36; Tsukada et al., 2006), JMJD1A/JHDM2A, which demethylates mono- and dimethyl H3-K9 (Yamane et al., 2006), and JMJD2A, which demethylates trimethyl H3-K9/K36 (Whetstine et al., 2006). These enzymes belong to a large family of proteins that contain a conserved JmjC domain, which has been shown to be critical for demethylation (Trewick et al., 2005; Tsukada et al., 2006). Together, these recent discoveries suggest an important but still mostly unknown role exerted by demethylation events in regulated gene transcription.

Here we report an unexpectedly widespread recruitment of the histone demethylase LSD1 to active promoters, including most ERα gene targets in MCF7 cells, and we find that LSD1 is needed for activation and is required to oppose the functions of three HMTs: RIZ1, ESET, and Eu-HMTase1. For gene transcription, these HMTs exert an inhibitory gatekeeper function, which is required to prevent recruitment of unliganded nuclear receptors and constitutive activation. Intriguingly, a similar molecular strategy is also employed for ERα gene targets that do not recruit LSD1 but utilize distinct combinations of HMTs and HDMs to provide promoter specificity to this HMT/HDM code. A similar code appears to exist also for a number of other signal-dependent transcription programs. In summary, we propose that a gatekeeper strategy imposed by a selective use of HMTs and HDMs is widely exploited in mammalian biology to determine the appropriate physiological amplitude of regulated gene transcription as orchestrated by nuclear receptors and other classes of DNA-binding transcription factors.

RESULTS

A Broad Genome-Wide LSD1 Promoter-Binding Program Detected by ChIP-DSL

Based on our interest in the role of corepressor complexes in gene transcription regulation, we applied a new genome-wide chromatin immunoprecipitation (ChIP) assay based on DSL (*D*NA *s*election and *l*igation; details in Supplemental Experimental Procedures and in Kwon et al., 2007) to several CoREST corepressor complex components, including the specific histone lysine demethylase LSD1. Recruitment of LSD1 in MCF7 cells that had been treated with 17 β-estradiol (E_2) was determined by the ChIP-DSL method using a specific α-LSD1 antibody (Figure S1) and an array containing 20,045 human-proximal promoters (Hu20K; Figure 1A). Statistical significance for promoter enrichment was based on the analysis of a genomic tiling component included on the array (Figure 1A, yellow dots). Based on data obtained from three independent biological replicates, we unexpectedly detected a total of 4212 LSD1-enriched ($LSD1^+$) promoters in these cells

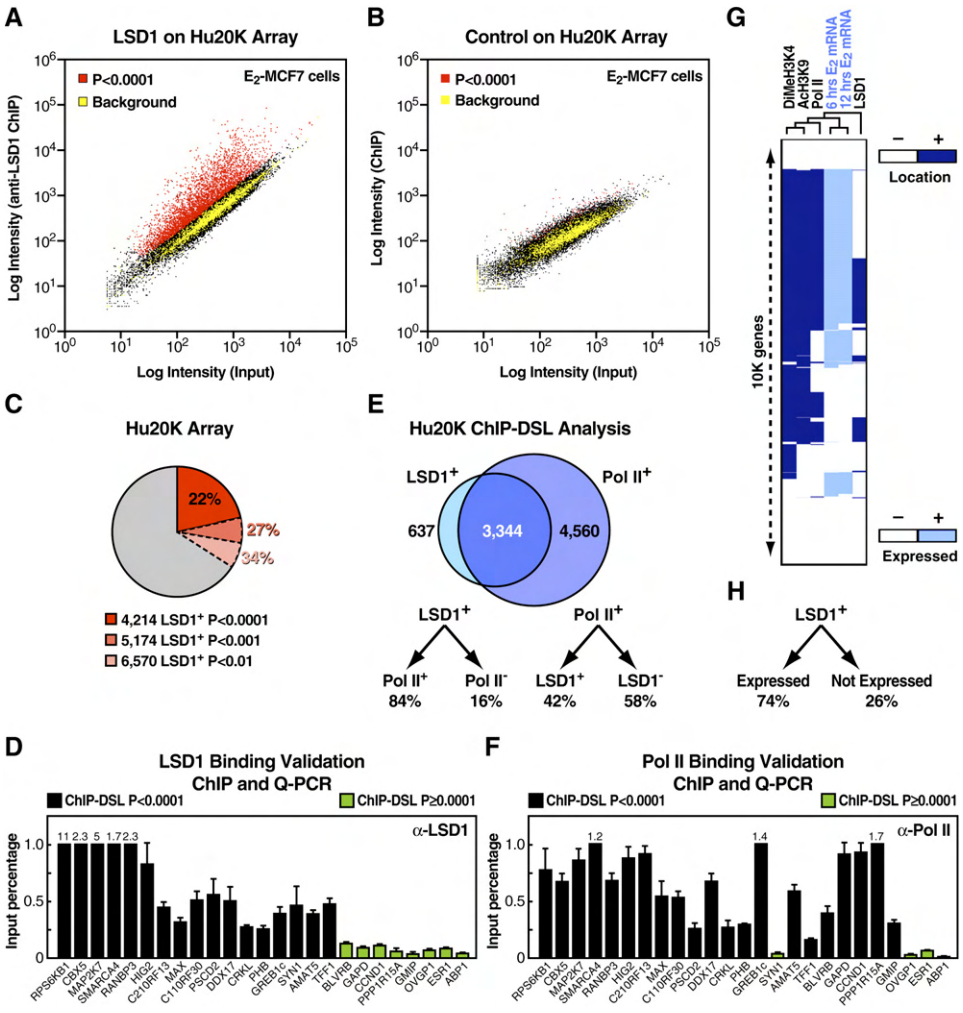

Figure 1. Genome-Wide Promoter Analysis Reveals the Association of LSD1 with a Broad Gene Activation Program

(A) Scatter plot of LSD1 recruitment to human gene promoters is shown. E_2-induced MCF7 cells were profiled on the Hu20K array by ChIP-DSL. Three biological replicates were used to derive enriched promoters at $p < 0.0001$ (red), shown in comparison with intergenic genomic sequences as negative controls (yellow). Weighted average is visualized on a single experiment scatter plot.

(B) Scatter plot of a nonrelevant antibody profiled on the Hu20K array by ChIP-DSL is shown. Two biological replicates were used.

(C) Number and percentage of $LSD1^+$ promoters on the Hu20K array at different statistical cutoffs are shown.

(D) ChIP/qPCR analysis of 17 $LSD1^+$ and 8 $LSD1^-$ randomly selected promoters in E_2-stimulated MCF7 cells is shown. The data are the average of three replicates, and error bars represent ± standard error mean.

(E) Venn diagram of $LSD1^+$ and Pol II^+ promoters obtained by ChIP-DSL is shown. Only promoters with reliable signal intensities in both profiling experiments were included in the comparison.

(F) ChIP/qPCR analysis of Pol II recruitment on selected $LSD1^+$ and $LSD1^-$ promoters in E_2-stimulated MCF7 cells is shown. The data are the average of three replicates, and error bars represent ± standard error mean.

(G) Correlation of gene expression (light blue) with promoter occupancy (dark blue) is shown and includes histone modification marks (DiMeH3K4 and AcH3K9) as well as Pol II and LSD1, which were profiled in E_2-induced MCF7 cells. Only genes included in both promoter- and expression-profiling arrays and reliably scored in all measurements were used to construct the binary map by unsupervised hierarchical clustering analysis.

(H) $LSD1^+$ genes classified by mRNA expression status are shown.

(22% of the Hu20K array that showed reliable signals) even at a highly stringent selection criteria ($p < 0.0001$), while 118 promoters were detected with an irrelevant antibody using the same criteria (Figures 1A and 1B, red dots). The number of $LSD1^+$ promoters would increase to 5174 promoters (27% of the array) or 6570 promoters (34% of the array) if the selection criteria were relaxed to a more conventional stringency ($p < 0.001$ and $p < 0.01$, respectively; Figure 1C). These results were validated by conventional ChIP assay and quantitative PCR (qPCR) analysis on randomly selected $LSD1^+$ promoters at $p < 0.0001$ (Figure 1D).

Consistent with the predictive value of this assay, virtually all known LSD1-target promoters represented on the Hu20K array were detected as LSD1$^+$, including the repressed *SYN1* (Hakimi et al., 2002) and *SCN3A* (Shi et al., 2004) promoters (Figure S2). The multiple experimentally confirmed or computationally predicted NRSF/REST-target genes found as LSD1$^+$ (e.g., *CX36*, *DNM1*, *PAX4*, and *SYT2*; Andres et al., 1999; Ballas et al., 2001; Hakimi et al., 2002; Lunyak et al., 2002; Figure S3A) and the enrichment of NRSF/REST-predicted sites observed in LSD1$^+$ neuronal-specific genes (Figure S3B) are consistent with the model that, as a component of the CoREST complex, LSD1 is functionally important in negatively regulating neuronal NRSF/REST-dependent genes in non-neuronal cells (Andres et al., 1999; Ballas et al., 2001, 2005; Humphrey et al., 2001; You et al., 2001; Hakimi et al., 2002; Lunyak et al., 2002; Shi et al., 2003, 2004; Bruce et al., 2004).

Broad LSD1 Promoter Binding Is Primarily Associated with Gene Activation Events

In addition to the role of LSD1 in transcriptional repression (Figures S2 and S3), ~84% of LSD1$^+$ promoters were Pol II$^+$ and, quite unexpectedly, a full 42% of all Pol II$^+$ promoters in MCF7 cells were LSD1$^+$ (Figure 1E). These results were confirmed by conventional ChIP/qPCR assay (Figure 1F). Similarly, examination of histone marks was consistent with the preponderance of activated, rather than repressed, LSD1$^+$ promoters. Indeed, dimethylation of histone H3 lysine 4 (diMeH3K4), a substrate for LSD1 demethylation activity associated with LSD1 repressive functions (Shi et al., 2004, 2005; Lee et al., 2005), and acetylation of histone H3 lysine 9 (acH3K9), another well-established mark that correlates with gene activation, exhibited a considerable overlapping pattern with the presence of LSD1 on gene promoters (Figure 1G). These observations were corroborated by clustering ChIP-DSL and mRNA-profiling data sets obtained in MCF7 cells, thus observing ~74% LSD1$^+$ genes as expressed (Figures 1G and 1H). Taken together, our data suggest that the percentage of activated transcription units modulated by LSD1 (~74%–84%) is substantially larger than the percentage of activated transcription units (~16%–26%) included in the LSD1-repressed program (Figures 1E and 1H).

Recently, we identified 578 promoters that exhibit ERα binding in MCF7 cells (Kwon et al., 2007). Surprisingly, 58% of ERα-enriched (ERα$^+$) promoters also exhibited LSD1 recruitment (319 ERα$^+$/LSD1$^+$ promoters); these represented ~9% of all LSD1$^+$ promoters (Figure 2A). In addition, genomic tiling analysis using ChIP-DSL assay of two well-established ERα-regulated target genes, *TFF1/pS2* and *GREB1*, revealed LSD1- and ERα-corecruitment on presumed ERα-binding enhancer sites that were located ~10 Kb upstream of *TFF1/pS2* and ~40 Kb upstream of *GREB1* gene start sites (Figure 2B). This ChIP-DSL analysis also showed that an ERα coactivator, CBP, and an "activation" histone mark, AcH3K9, exhibited similar profiles (Figure 2B). For all factors tested, including LSD1, the occupancy on both distal and proximal promoter cognate DNA sites was shown to be stimulated by E_2 treatment as determined by ChIP/qPCR (Figure 2C). In addition, LSD1 recruitment on promoters of other LSD1$^+$/ERα$^+$-target genes was also stimulated by E_2 (Figure 2D). In agreement with these data, coimmunoprecipitation experiments in whole MCF7 cell extract using α-LSD1 antibody demonstrated a specific E_2-dependent interaction of LSD1 with endogenous ERα and CBP, and vice versa (Figures 2E and 2F).

Ligand-Dependent Induction of ERα-Regulated Targets Requires Functionally Active LSD1

We next assessed the functional requirement of LSD1 in gene activation of ERα-regulated genes and found that E_2-dependent induction of LSD1$^+$/ERα$^+$-target genes *(pS2, GREB1, CTSD,* and *MYC*) was essentially abolished in MCF7 cells in which *LSD1* was knocked down by siRNA (Figures 3A, 3B, and S4A). By contrast, analysis on an LSD1$^-$/ERα$^+$-target gene, *WISP2*, revealed that depletion of LSD1 had no effect on its ligand-dependent activation (Figure 3C).

To test the role of the amine oxidase activity of LSD1 on regulation of LSD1$^+$/ER$^+$ targets, a rescue experiment was performed by single-cell nuclear microinjection assay; in this experiment, wild-type human *LSD1* (*wt hLSD1*) or a predicted amine oxidase-inactive human form (*mut hLSD1*) was injected in Rat-1 cells that were simultaneously depleted of the endogenous rat LSD1 by microinjection of *rLSD1*siRNA. While WT hLSD1 rescued the E_2-induced activation of a *pS2* promoter *LacZ* reporter, the mut hLSD1 was unable to rescue the reporter gene activity (Figure 3D). Similarly, transient transfection of *wt hLSD1* but not *mut hLSD1* caused an E_2-dependent increase in the activity of an ERE-*Luc* reporter (Figure S5). Therefore, the amine oxidase activity, which was proven to be required for mediating LSD1-dependent H3-K4 and H3-K9 demethylation (Shi et al., 2004; Metzger et al., 2005), is also necessary for LSD1/ERα-dependent gene activation.

We next addressed any H3-K4 and H3-K9 demethylation event that might occur upon E_2 stimulation of LSD1/ERα-dependent gene targets. While ERα binding and acH3K9 were increased as a result of E_2 treatment (Figures 3E and 3F), we observed, by quantitative ChIP/qPCR on promoter and distal *pS2* and *GREB1* ERα-target sites, a consistent decrease in both diMeH3K4 and diMeH3K9 (Figures 3G and 3H). Importantly, no simultaneous increases in trimethylation status were observed on the same sites (Figures 3I and 3J). In concert, these findings suggest that H3-K9 and H3-K4 methylation dynamics are key components of the gene activation program mediated by LSD1.

Mechanisms Underlying the LSD1-Dependent Gene Activation

To understand the mechanism underlying the functional requirement for LSD1 in E_2-dependent induction, we examined the impact of removing various H3-K9 and

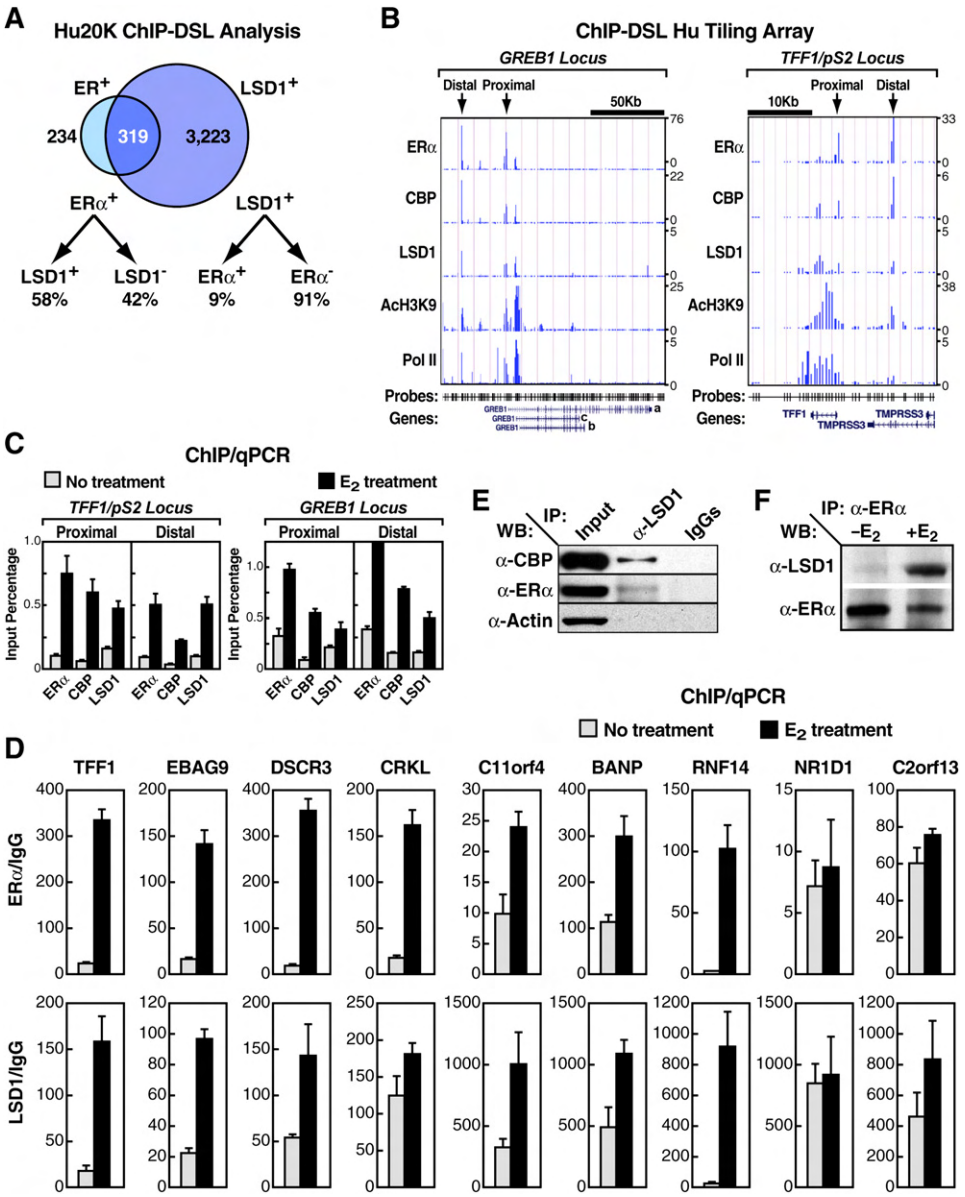

Figure 2. LSD1 Associates with Most ERα-Promoter Targets in MCF7 Cells

(A) Venn diagram of LSD1$^+$ and ERα$^+$ promoters as obtained by ChIP-DSL is shown. Only promoters with reliable signal intensities in both profiling experiments were included in the comparison.

(B) ChIP-DSL tiling array analysis on *GREB1* (left panel) and *TFF1/pS2* (right panel) loci of ERα, CBP, LSD1, AcH3K9, and Pol II occupancy in E_2-stimulated MCF7 cells is shown. Binding profiles represent ChIP versus input DNA intensity ratios (right). The most enriched ERα-binding sites (proximal and distal) are indicated by arrows. DSL probe location and RefSeq gene annotation are indicated in the bottom.

(C) ChIP/qPCR recruitment analysis of ERα, CBP, and LSD1 on proximal and distal ERα-binding sites upon E_2 treatment in MCF7 cells is shown. The data are the average of three replicates, and error bars represent ± standard error mean.

(D) ChIP/qPCR analysis of LSD1 and ERα recruitment on additional LSD1$^+$/ERα$^+$ promoters detected by ChIP-DSL assay upon E_2 treatment in MCF7 cells is shown. The data are the average of three replicates, and error bars represent ± standard error mean.

(E) shows coimmunoprecipitation analysis of ERα and CBP by anti-LSD1 antibody in cell extracts obtained from E_2-stimulated MCF7 cells. 5% input is shown.

(F) Coimmunoprecipitation analysis of LSD1 by anti-ERα antibody in cell extracts obtained from unstimulated and E_2-stimulated MCF7 cells is shown.

H3-K4 HMTs (Martin and Zhang, 2005) by siRNA on E_2-dependent induction. Depletion of known H3-K4 HMTs, including MLL1, had little effect on basal *pS2* activity (data not shown). However, depletion of several specific H3-K9 HMTs, including RIZ1, ESET, and Eu-HMTase1, but not G9a or Suv39h1/h2, derepressed the *pS2*

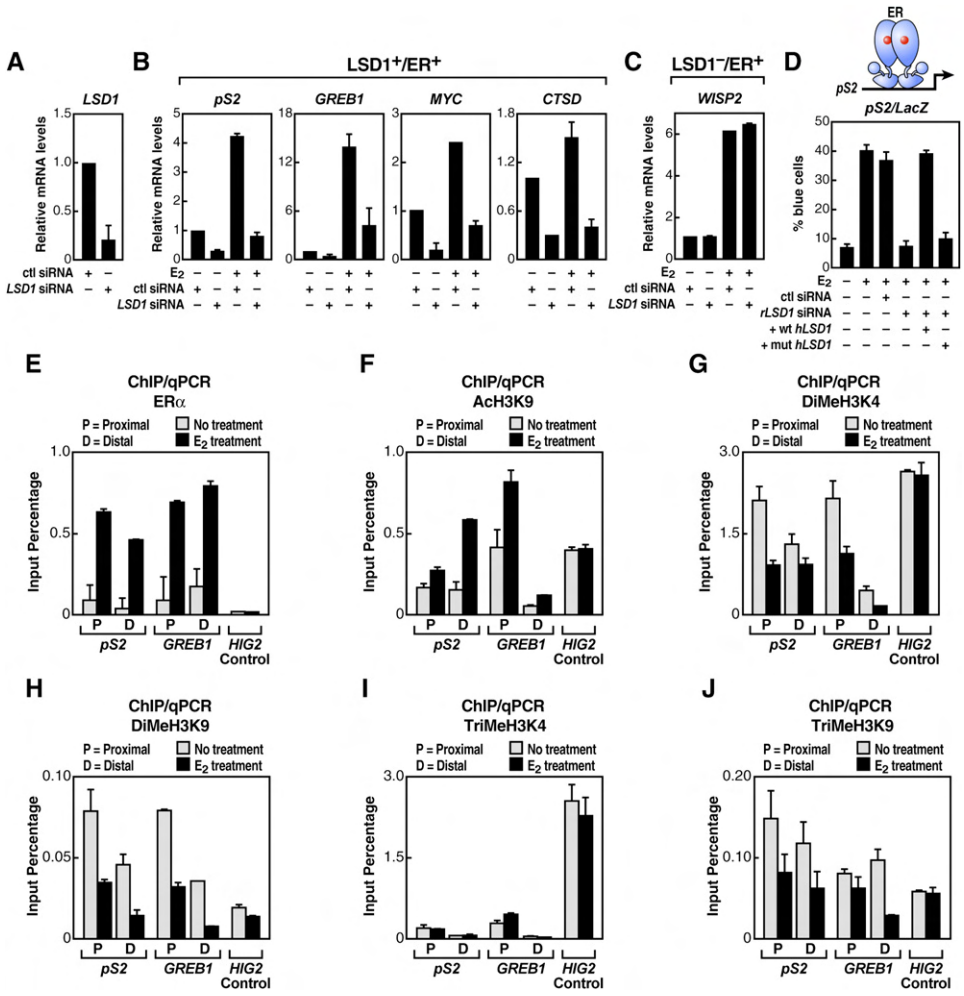

Figure 3. LSD1 Regulates E_2-Dependent Gene Transcription While Both H3-K4 and H3-K9 Demethylation Events Are Observed

(A) Real-time qPCR (RT-qPCR) analysis is shown to document efficiency of *LSD1* siRNA to diminish endogenous *LSD1*.

(B) RT-qPCR analysis of several endogenous LSD1$^+$/ERα^+-target genes upon *LSD1* siRNA transfection is shown.

(C) RT-qPCR analysis of an endogenous LSD1$^-$/ERα^+-target gene upon LSD1 depletion by siRNA is shown. In (A)–(C) *LSD1* siRNA was delivered by transient transfection in MCF7 cells, and β-actin mRNA expression levels as well as cell transfection efficiency were used for normalization.

(D) Functional rescue analysis of the wild-type (wt) and the amine oxidase mutant (mut) human LSD1 form (hLSD1) in Rat-1 cells is shown. Endogenous rat LSD1 (rLSD1) expression was depleted by specific rat *LSD1* siRNA, and ectopic wt hLSD1 and mut hLSD1 overexpression was accomplished by expression plasmids. Reporter plasmid and *LSD1* siRNA were delivered by single-cell nuclear microinjection in Rat-1 cells.

(E–J) Panels show ChIP/qPCR occupancy analysis of ERα (E), acH3-K9 (F), diMeH3-K4 (G), diMeH3-K9 (H), TriMeH3-K4 (I), and TriMeH3-K9 (J) on ERα-binding sites in vehicle- or E_2-treated MCF7 cells. ERα proximal or distal binding sites on *pS2* and *GREB1* genomic loci were examined; *HIG2* promoter was included as control. The data in (A)–(J) are the average of three replicates, and error bars represent ± standard error mean.

promoter; this resulted in full activation in the absence of ligand (Figure 4A), while no significant effects were observed after depletion of these three enzymes in the presence of ligand (Figure 4B). Thus, in the absence of these specific H3-K9 HMTs, regulated activation of *pS2* gene now became LSD1 independent (Figure 4C). Consistent with these observations, mRNA levels of endogenous *pS2* and *GREB1* genes were upregulated upon siRNA-mediated RIZ1/ESET depletion even in the absence of ligand without any effect on response observed in the presence of ligand (Figures 4D, 4E, and S4B). To demonstrate the specificity in knockdown experiments, we observed at least a partial (~70%) functional rescue of the repressive ESET activity in HeLa cells in which the endogenous human ESET (hESET) was depleted by specific *hESET* siRNA, and the mouse form of this HMT (mESET) was overexpressed (Figure 4F). In addition, the binding of ESET and RIZ1 on *pS2* and *GREB1* promoters was found to be decreased upon E_2 stimulation as quantified by ChIP/qPCR (Figure 4G). Together, these data suggest that a critical function of LSD1 is to reverse the inhibitory effect of several H3-K9 HMTs: RIZ1, ESET, and Eu-HMTase1, the presence of which dictates LSD1 dependency for gene activation.

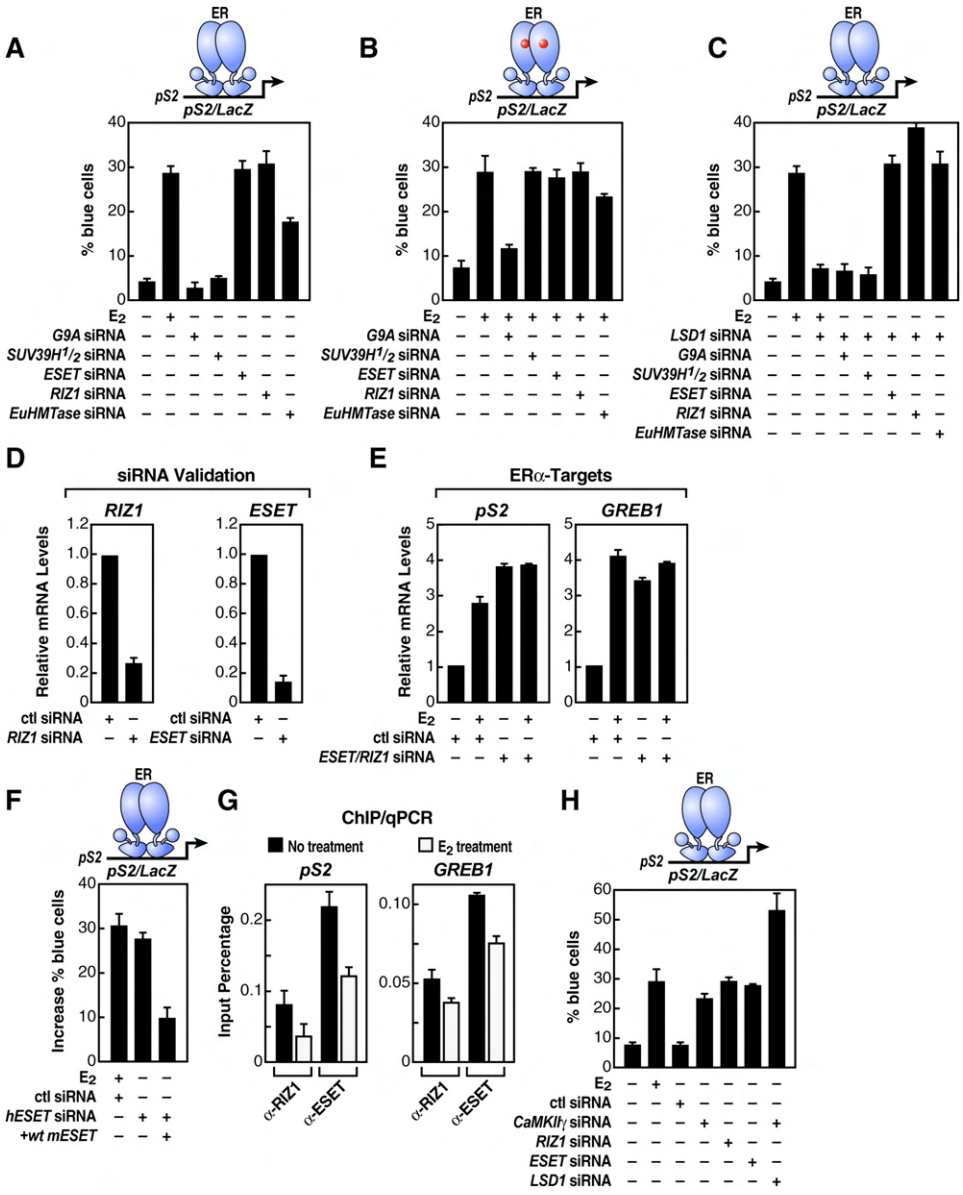

Figure 4. Specific H3-K9 HMTs Function as Inhibitory Gatekeepers and Dictate LSD1 Dependency for ERα-Regulated Gene Activation

(A and B) Effect of different siRNAs to specific H3-K9 HMTs on *pS2* promoter-*LacZ* reporter activity was analyzed by single-cell microinjection assay in the absence (A) and presence (B) of ligand (E_2).

(C) Functional LSD1 dependency of *pS2* promoter-*LacZ* reporter activity was tested after depletion of specific H3-K9 HMTs.

(D) RT-qPCR analysis was performed to document efficiency of *H3-K9 HMTs* siRNAs to diminish endogenous RIZ1 and ESET.

(E) RT-qPCR analysis of endogenous ERα-target genes upon specific *H3-K9 HMTs* siRNA transfection is shown.

(F) Functional rescue analysis of a mouse ESET form in HeLa cells is shown. Endogenous human ESET (hESET) expression was abolished by specific mouse *ESET* siRNA, and ectopic mouse ESET (mESET) overexpression was accomplished by expression plasmid.

(G) Panel shows ChIP/qPCR recruitment analysis of RIZ1 and ESET on endogenous ERα-target promoters upon E_2 stimulation in MCF7 cells.

(H) Functional analysis of CaMKIIγ in absence of ligand (E_2) on a *pS2* promoter-*LacZ* reporter gene is shown. For experiments (A)–(C), (F), and (H) reporter plasmid and siRNAs were delivered by single-cell nuclear microinjection in MCF7 and HeLa (in F) cells. In (D) and (E) siRNAs were delivered by transient transfection in MCF7 cells, and β-actin expression levels and cell transfection efficiency were used for normalization. The data in (A)–(H) are the average of three replicates, and error bars represent ± standard error mean.

The requirement for more than one H3-K9 HMT implies that, at least for some genes, these enzymes are not fully functionally redundant; this has also been recently shown for similar nonredundant HDM activities (Yamane et al., 2006). Interestingly, at least one of the specific H3-K9 HMTs, *RIZ1,* has been proposed to act

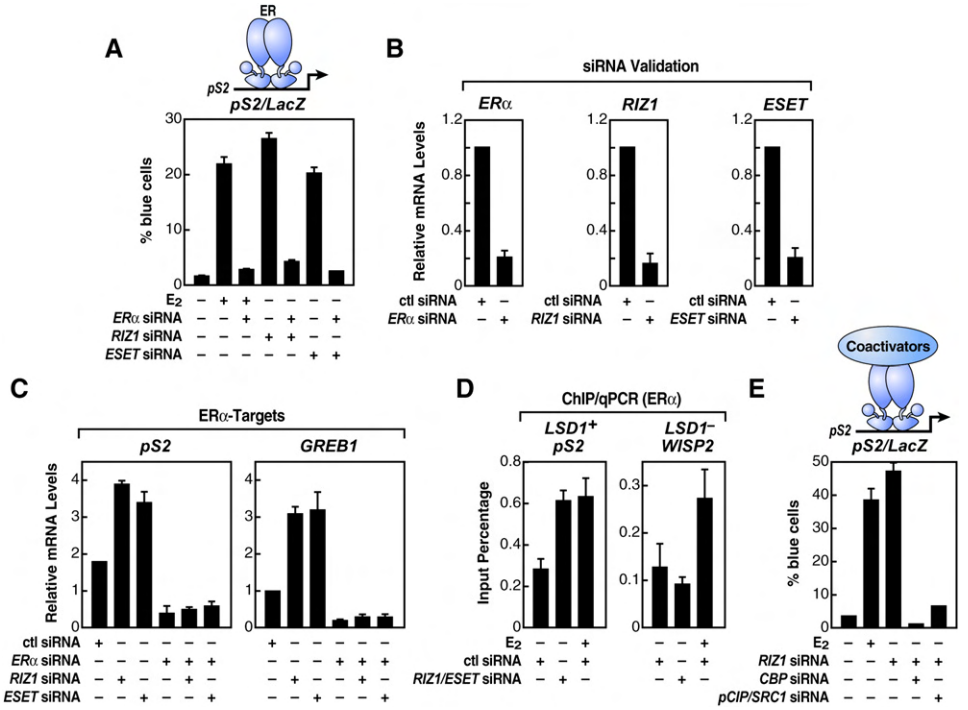

Figure 5. H3-K9 HMTs Impose Ligand Dependency to the E_2-Dependent Signaling Pathway

(A) ERα dependency of the *pS2* promoter-*LacZ* reporter activity is analyzed in absence of specific H3-K9 HMTs.

(B) RT-qPCR analysis was performed to document efficiency of *ERα*, *RIZ1,* and *ESET* siRNAs to diminish endogenous ERα, RIZ1, and ESET.

(C) RT-qPCR analysis of endogenous ERα-target genes after removing H3-K9 HMTs and ERα by siRNA is shown.

(D) ChIP/qPCR recruitment analysis of ERα in cells transfected with specific *H3-K9 HMTs* siRNAs in absence or presence of ligand (E_2) is shown.

(E) Analysis of *pS2* promoter-*LacZ* reporter activity upon removing the ERα-associated coactivators CBP, pCIP, and SRC1 in cells depleted of the H3-K9 HMT RIZ1. For experiments (A) and (E), reporter plasmid and siRNAs were delivered by single-cell nuclear microinjection in MCF7 cells. In (B)–(D), siRNAs were delivered by transient transfection in MCF7 cells, and β-actin mRNA expression levels and cell transfection efficiency were used for normalization. The data in (A)–(E) are the average of three replicates, and error bars represent ± standard error mean.

as a tumor-suppressor gene that is silenced in many highly metastatic breast cancer cell lines (He et al., 1998; Du et al., 2001), and *RIZ1* siRNA-depleted MCF7 cells showed increased levels of cell proliferation (Carling et al., 2004; Gazzerro et al., 2006). Indeed, cell proliferation and cell-cycle control were the top two ER^+ gene ontology-enriched terms that were revealed by ChIP-DSL analysis of ERα in MCF7 cells (Kwon et al., 2007), suggesting a global upregulation of ER^+ gene targets in the absence of RIZ1.

Because of the connection between specific protein kinases, including CaMKs, and the actions of repressive HMTs (Ishitani et al., 2003; Zhang, et al., 2002; Kurahashi et al., 2005), we tested the possibility that CaMKII might modulate H3-K9 HMTs and, hence, LSD1- and E_2-dependent gene activation events. Interestingly, specific *CAMKIIγ* siRNA resulted in full, LSD1-independent activation of the *pS2*-reporter gene (Figure 4H), which strongly suggests a link between specific signaling pathways and an HMT-dependent imposition of inhibitory histone marks.

A Balance between Histone Lysine Methylation and Demethylation Events Determines Ligand Dependency of ERα Targets

The observation of "ligand-independent" gene activation in the absence of specific H3-K9 HMTs raised the question of whether ERα is still critically required for activation in the absence of these enzymes and LSD1. To investigate this issue, siRNA depletion of ERα, along with RIZ1 or ESET, was performed in MCF7 cells and revealed that, in the absence of ERα, *pS2* promoter activation was abolished (Figure 5A). These results, which are fully in concert with the observation that unliganded ERα can still bind to and even recycle at low levels on the *pS2* promoter (Metivier et al., 2004; Perissi et al., 2004), were confirmed on endogenous ERα-target genes (Figures 5B and 5C).

ChIP analysis of RIZ1/ESET siRNA-depleted cells showed an increased recruitment of ERα, in absence of E_2, to a level equivalent to that observed with E_2, thus supporting the hypothesis that one critical function of H3-K9 HMTs is to inhibit recruitment of unliganded ERα to its target genes (Figure 5D). Furthermore, well-established ERα

coactivators, including CBP and SRC1-p/CIP, were required for this LSD1- and ligand-independent activation (Figure 5E). In contrast, the depletion of these same HMTs caused neither activation (data not shown) nor increased ERα binding on the LSD1-independent ERα target *WISP2* promoter (Figure 5D), suggesting a role of other combinations of HMTs/HDMs in ERα regulation of LSD1-independent genes (see below).

Based on the broad recruitment of LSD1 on proximal promoters (Figure 1C), we determined the LSD1 dependency of several additional regulated gene activation programs. *LSD1* siRNA microinjection experiments revealed a similar role of LSD1 in activation of AR, as previously shown (Metzger et al., 2005; Yamane et al., 2006). In addition, these experiments revealed a role of LSD1 in activation of NFκB-, AP-1-, and βRAR-regulated promoters, but not in the case of CREB-regulated gene targets (Figure S6). These data suggest that LSD1 is selectively involved in induction of a subset of regulated transcription units but is not universally required for gene activation. Analogous to ERα, RIZ1 and ESET appeared to exert a similar role in AR-dependent activation of the *PSA* gene (Figures 6A and S7A) as suggested by the partial derepression of *PSA* expression levels and the LSD1-independent activation of the same gene, which were again produced in the absence of specific H3-K9 HMTs. Similarly, depletion of RIZ1/ESET resulted in activation of NFκB-target genes in absence of signal, and this activation also exhibited dependency on the presence of NFκB factors (Figures S7B–S7D). Based on these data, we propose that a major biological role for the opposing actions of RIZ1/ESET/Eu-HMTase1 and LSD1 is to confer high amplitude of ligand response or signal dependency of gene activation on cohorts of regulated transcription units.

Promoter-Specific HMT/HDM "Code" for Regulated Transcription Units

Because *pS2* and *GREB1* require LSD1 for activation, while *WISP2* does not, a key question is whether entirely distinct molecular strategies are used for $LSD1^+$- and $LSD1^-$-regulated transcription units or whether distinct HMTs and HDMs are required to mediate a similar regulatory mechanism. As a new AR-specific H3-K9 HDM, JMJD1A/JHDM2A, has recently been reported (Yamane et al., 2006), we examined whether this enzyme exerts a similar role on $LSD1^+$/ERα-regulated targets and, alternatively, whether this factor might substitute for LSD1 on $LSD1^-$/$ER\alpha^+$ gene targets. The specific depletion of JMJD1A by siRNA in LNCaP cells confirmed that this factor was required for ligand-dependent activation of AR-target genes (Figures 6B and 6C), as previously shown (Yamane et al., 2006). Because LSD1 is also required (Figure 6A; Metzger et al., 2005), these AR-target genes demonstrate that at least two HDMs are required for full activation (Yamane et al., 2006). In contrast, ERα-dependent activation of *pS2* and *GREB1* gene targets depends on LSD1 but not on JMJD1A when similar experiments are performed in MCF7 cells (Figures 6D–6F). Interestingly, however, an unexpected activation of the *pS2* promoter was observed in the absence of ligand (Figures 6D–6F). Based on these findings, we next explored whether an $LSD1^-$/$ER\alpha^+$ gene target, *WISP2*, might use JMJD1A and HMTs other than ESET/RIZ1/EuHMTase to mediate a strategy of regulation similar to that utilized by LSD1 and these enzymes for *pS2*. Interestingly, while *LSD1* siRNA did not inhibit ligand-dependent activation (Figure 3C), *JMJD1A* siRNA blocked E_2-dependent stimulation of *WISP2* (Figure 6G), and *G9A* but not *SUV39H1*/*H2* siRNAs permitted a partial activation (~50%) in the absence of ligand (Figure 6H). Together, these results provide compelling evidence for a widely used, promoter-specific code in regulated transcription, with different cohorts of targets requiring distinct combinations of HMTs and HDMs. It is interesting to note that for some genes, loss of HMTs alone causes a level of activation similar to that observed in response to ligand, while for others, only a partial activation, with further enhancement when ligand is present, is observed.

DISCUSSION

Here we report an unexpected mechanism that proves to be physiologically utilized to prevent constitutive gene activation by unliganded nuclear receptors and other classes of signal-dependent DNA-binding transcription factors. We have identified three specific H3-K9 HMTs that potentially act as gatekeepers for the ~58% of ERα gene targets that are $LSD1^+$ and that maintain a repressive status that precludes unliganded nuclear receptors, such as ERα, from effective binding and from functioning as ligand-independent, constitutive activators (Figure 6I). Indeed, transcriptional activation by unliganded nuclear receptors has been reported on removal of corepressors and on activation by specific signaling pathways (Culig et al., 1994; Nazareth and Weigel, 1996; Zwijsen et al., 1997; Rogatsky et al., 1999; Ueda et al., 2002; Ogawa et al., 2004; Kim et al., 2005).

Our data support the model that inhibitory HMTs prevent unregulated activation of these targets, at least in part, by blocking the binding of unliganded ERα to its cognate DNA sites. The binding of liganded ERα on these regulated promoters thereafter permits the recruitment of the LSD1 enzymatic activity that is required to dismiss the inhibitory marks and, potentially, methyl marks of other nonhistone substrates (I.G.-B. and M.G.R., unpublished results), thus permitting other ERα-recruited coactivators to initiate the regulated transcriptional response (Figure 6I). Thus, exchange of inhibitory histone lysine methyl marks based on the actions of specific, and not fully redundant, H3-K9 HMTs as well as the actions of critical HDMs are required to permit a ligand-dependent response. Therefore, while REST/CoREST-dependent genes clearly employ LSD1 in repression for at least a cohort of the REST-dependent program (Shi et al., 2004; Figures S2 and S3), our results have also revealed that LSD1, most likely associated with the MLL1 complex (Nakamura

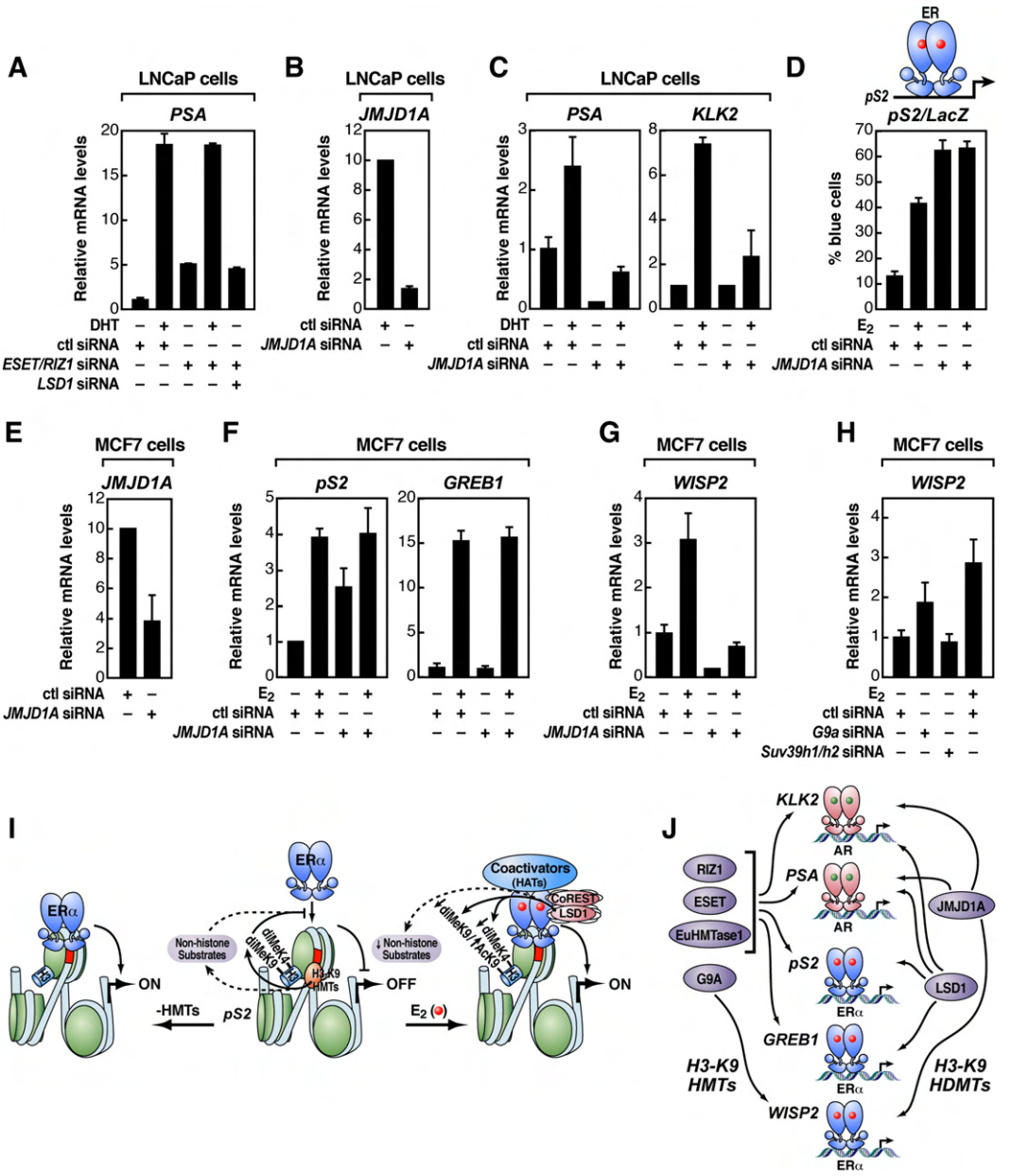

Figure 6. A Promoter-Specific HMT/HDM "Code" for Regulated Transcription Units

(A) RT-qPCR analysis of an endogenous LSD1^{+}/AR^{+} gene target upon RIZ1/ESET and/or LSD1 depletion by siRNA in LNCaP cells is shown.

(B) RT-qPCR analysis to document the efficiency of *JMJD1A* siRNA to diminish endogenous *JMJD1A* in LNCaP cells is shown.

(C) RT-qPCR analysis of endogenous AR-target genes upon JMJD1A depletion by siRNA is shown.

(D) Functional analysis of the *pS2* promoter-*LacZ* reporter gene activity after removing JMJD1A by microinjection of siRNA in MCF7 cells is shown.

(E) RT-qPCR analysis was performed to document efficiency of *JMJD1A* siRNA to diminish endogenous JMJD1A in MCF7 cells.

(F) RT-qPCR gene expression analysis of endogenous LSD1^{+}/ERα^{+}-target genes upon JMJD1A depletion by siRNA is shown.

(G) RT-qPCR analysis of an endogenous LSD1^{-}/ERα^{+}-target gene upon JMJD1A depletion by siRNA is shown.

(H) RT-qPCR gene expression analysis of an endogenous LSD1^{-}/ERα^{+}-target gene upon G9a or Suv39h1/h2 depletion by siRNA is shown.

(I) Model of H3-K9 HMT requirement to inhibit constitutive ERα activation by blocking binding of the unliganded nuclear receptor to its cognate DNA sites is shown; HDMs, as LSD1, are required to demethylate H3-K9 HMTs substrates to permit activation by liganded ERα (see text for details).

(J) Gene-specific use of HMT/HDMs to define regulated gene activation programs (see text for details) is shown. In all these experiments, siRNA was delivered by transient transfection in LNCaP (A)–(C) or MCF7 cells (D)–(H), and β-actin expression levels and cell transfection efficiency were used for normalization. The data in (A)–(H) are the average of three replicates ± standard error of the mean.

et al., 2002; Guenther et al., 2005), is required for a surprisingly broad range of regulated gene activation events. We further suggest that the HMT/HDM cycle provides a licensing mechanism for signal-dependent gene activation of other classes of regulated transcription factors that are required for normal development and homeostasis.

Gene-Specific Use of HMT/HDMs to Define Regulated Gene Activation Programs

Because LSD1 is recruited only to a subset of ERα-target genes, we were keen to initially examine the possibility that a similar inhibitory histone code and degree of HMT/HDM usage might act on genes even when LSD1 is not recruited to the promoter. Recently it has been reported that the mono- and dimethyl H3-K9 HDM, JMJD1A/JHDM2A, contributes with LSD1 to the hormone-induced H3-K9 demethylation of the *PSA* promoter/enhancer (Metzger et al., 2005; Yamane et al., 2006), and that an additional H3-K9 demethylase, JMJD2A, displays specific trimethyl H3-K9/K36 activity (Whetstine et al., 2006). Therefore, analogous to histone deacetylase enzymes (HDACs), several H3-K9 HMTs and HDMs could be present in the same multiprotein complex (McKinsey et al., 2001; Yang and Seto, 2003; Ogawa et al., 2002; Shi et al., 2003) and might act to regulate gene expression in a nonredundant fashion involving additional, distinct substrates.

Here, we have shown that distinct HMTs and HDMs are utilized in a promoter-specific fashion (Figure 6J). For example, the LSD1$^-$/ERα$^+$ *WISP2*-target gene requires JMJD1A for ligand-dependent activation, in which G9a is selectively required to prevent ligand-independent constitutive activation. However, LSD1 but not JMJD1A is necessary for E_2-dependent activation of the ERα$^+$-target genes *pS2* and *GREB1*, while both LSD1 and JMJD1A are required for DHT activation of the *KLK2* and *PSA* genes. It is therefore tempting to speculate that there will be multiple cohorts of target genes for each nuclear receptor that depend on a similar strategy to prevent constitutive activation, but each will utilize unique combinations of HMTs and HDMs to license ligand-regulated expression.

While we would predict the existence of additional substrates other than H3-K9, for H3-K9 HMTs, and H3-K4/K9, for LSD1, we propose that ultimately the inhibitory gatekeeper strategy that is imposed by different HMTs blocks ligand-independent receptor binding and constitutive gene activation. This gatekeeper function is biologically overcome on binding of liganded receptors and specific H3-K9 HDMs. The requirement for three different HMTs, RIZ1/PRDM5 (Steele-Perkins et al., 2001), ESET/SETDB1 (Yang et al., 2002; Schultz et al., 2002; Wang et al., 2003), and Eu-HMTase1/GLP (Ogawa et al., 2002), for at least some ERα-target genes, suggests that these SET-domain enzymes are not fully functionally redundant (Schultz et al., 2002; Wang et al., 2003), as seems to similarly occur for at least two HDMs, LSD1 and JMJD1A/JHDM2A, for the regulation of some AR-target genes (Metzger et al., 2005; Yamane et al., 2006). As this represents an analogous strategy involving different HMTs and HDMs, the link between histone marks in preventing constitutive activation of physiologically regulated genes appears to be a general strategy.

H3-K4 and H3-K9 Demethylation in E_2-Dependent Gene Activation

At least two different dimethylated lysines have been proposed to be demethylated by LSD1: the well-documented H3-K4 (Shi et al., 2004, 2005; Lee et al., 2005) and H3-K9 (Metzger et al., 2005). LSD1 was found by ChIP-DSL to be mainly associated with the presence of dimethyl H3-K4 on the genome (Figure 1G). Indeed, although a certain degree of H3-K4 demethylase activity was observed upon activation of ERα targets, high remaining levels of this mark were still observed after LSD1 binding and activation, which would agree with the high overlapping pattern detected (Figure 3G). In contrast, it is proposed that this mark is fully removed by LSD1 on targets that are repressed by NRSF, which is supposed to confer gene repression (Shi et al., 2004). Thus, LSD1 might switch activity depending on the promoter, which is consistent with the ability of CoREST and BHC80 to modify LSD1 enzymatic activity in vitro (Shi et al., 2005; Lee et al., 2005), therefore displaying H3-K4 demethylase on NRSF targets and H3-K9 demethylase on AR, ERα, and other factors' target promoters. Also, one of the two reported LSD1 activities might be selectively inactivated by the surrounding histone marks present around the promoter. Indeed, specific histone marks, such as AcH3-K9, inhibit H3-K4 demethylation in vitro (Forneris et al., 2005, 2006). Of course, both activities could be also equally utilized on all promoters, with the final levels of histone methylation being indeed determined by the predemethylation levels.

The regulatory roles of these two histone methylation marks on gene activation of euchromatic promoters have been the subject of intense investigation during the past few years. For example, methylation of H3-K9 has been described to decrease on a subset of inflammatory genes upon LPS stimulation (Saccani and Natoli, 2002). Similarly, an inverse relationship between H3-K9 methylation and NFκB p65 binding on the IL-1β promoter has been observed upon LPS stimulation in THP-1 cells (Chan et al., 2005). Dimethylation of H3-K9 but also H3-K4 has been described to decrease when T_3R shifted from an unliganded state to a liganded, active state (Li et al., 2002), and a dramatic decrease in H3-K4 occurred on a tandem array of the mouse mammary tumor virus (MMTV) promoter upon induction by glucocorticoids (Ma et al., 2001). Finally, a decline in both dimethyl H3-K4 (Kim et al., 2003) and dimethyl H3-K9 (Metzger et al., 2005) has been reported on the *PSA* promoter upon DHT induction, and here we show that both histone marks are simultaneously decreased upon ERα/E_2-dependent activation of several promoters. Thus, dynamic changes of H3-K4 and H3-K9 methylation seem to occur in a diversity of programs.

In conclusion, an important aspect of the serial post-translational methylation modifications of histones and

transcription factors, which constitute a component of the hypothesized "histone code," is to impose signal/ligand dependency to regulated DNA-binding transcription factors by modulatory binding of unliganded receptors and other classes of transcription factors, thereby preventing constitutive activation of cohorts of target genes.

EXPERIMENTAL PROCEDURES

Materials and Reagents

Anti-LSD1 antibody was generated in guinea pigs against bacterially expressed N′-terminal region (13–219 aa) of the human LSD1. Anti-CoREST antibody was a generous gift from Dr. Gail Mandel. Commercially available antibodies and siRNAs are documented in Supplemental Data.

ChIP Assay

MCF7 breast cancer cells were hormone deprived for 4 days and then treated for 1 hr with 100 nM 17β-estradiol (E_2). ChIP assay was conducted as previously described (Zhu et al., 2006). Further details and list of primers used for PCR validation are available in Supplemental Data.

ChIP-DSL Assay

Genome-wide promoter location analysis by ChIP-DSL on Hu20K and HuTiling arrays was performed and analyzed as described in Supplemental Experimental Procedures and in Kwon et al., 2007. Biological triplicates were performed in each case.

Single-Cell Nuclear Microinjection Assay

The small interfering RNAs (siRNAs) were delivered into cells by single-cell microinjection assay as previously described (Kamei et al., 1996; Heinzel et al., 1997). Details and siRNA information are available in Supplemental Experimental Procedures.

Transient Transfection Assay

Transient transfection assay in MCF7 cells was performed using Lipofectamin 2000, following manufacturer's protocol (Invitrogen).

RNA Isolation, Real-Time PCR, and RNA Profiling Analysis

RNA was extracted by RNeasy Assay (Qiagen), and cDNA was obtained by SuperScript First-Strand Synthesis System (Invitrogen) following both manufacturers' protocols. RNA profiling was performed as described by Kwon et al., 2007. Further details and list of primers used for PCR amplification are available in Supplemental Experimental Procedures.

Immunoprecipitation

Immunoprecipitation assay was conducted as previously described (Bannister and Kouzarides., 1996). Details are available in Supplemental Experimental Procedures.

Real-Time qPCR

Semi-qPCR was carried out according to the Mx3000P Real-Time PCR Systems manual and the Brilliant qPCR reagent kit (Stratagene). Details are available in Supplemental Experimental Procedures.

Supplemental Data

Supplemental Data include seven figures, Supplemental Experimental Procedures, and Supplemental References and can be found with this article online at http://www.cell.com/cgi/content/full/128/3/505/DC1/.

ACKNOWLEDGMENTS

We thank colleagues in our laboratories, especially C.A. Nelson and A. Krones for technical assistance. We also thank Dr. V. Perissi for critical reading of the manuscript and Dr. A. Almenar-Queralt for generous help; Dr. A. Aggarwal for advice on prediction of essential enzymatic LSD1 residues; Dr. G. Mandel for kindly providing CoREST antibodies; A. Gonzalez (Santa Cruz Technologies) for advice on reagents; and J. Hightower and M. Fisher for figure and manuscript preparation. We especially acknowledge Ron Margolis and Phil Smith (NIDDK) and Laura Mamounas (NINDS) for their critical support in permitting development of this technology. M.G.R. is an investigator with Howard Hughes Medical Institute. This work is supported by CA97134, DK39949, and NS34934 NIH grants, Vitamin Cases Consumer Settlement Fund, and Army Concept Award BC0464 to M.G.R.; DK063491 and CA52599 NIH grants to C.K.G.; and CA114184 and HG003119 NIH grants to X.-D.F.

Received: April 13, 2006
Revised: July 27, 2006
Accepted: December 4, 2006
Published: February 8, 2007

REFERENCES

Andres, M.E., Burger, C., Peral-Rubio, M.J., Battaglioli, E., Anderson, M.E., Grimes, J., Dallman, J., Ballas, N., and Mandel, G. (1999). CoREST: a functional corepressor required for regulation of neural-specific gene expression. Proc. Natl. Acad. Sci. USA *96*, 9873–9878.

Ballas, N., Battaglioli, E., Atouf, F., Andres, M.E., Chenoweth, J., Anderson, M.E., Burger, C., Moniwa, M., Davie, J.R., Bowers, W.J., et al. (2001). Regulation of neuronal traits by a novel transcriptional complex. Neuron *31*, 353–365.

Ballas, N., Grunseich, C., Lu, D.D., Speh, J.C., and Mandel, G. (2005). REST and its corepressors mediate plasticity of neuronal gene chromatin throughout neurogenesis. Cell *121*, 645–657.

Bannister, A.J., and Kouzarides, T. (1996). The CBP co-activator is a histone acetyltransferase. Nature *384*, 641–643.

Bruce, A.W., Donaldson, I.J., Wood, I.C., Yerbury, S.A., Sadowski, M.I., Chapman, M., Gottgens, B., and Buckley, N.J. (2004). Genome-wide analysis of repressor element 1 silencing transcription factor/neuron-restrictive silencing factor (REST/NRSF) target genes. Proc. Natl. Acad. Sci. USA *101*, 10458–10463.

Carling, T., Kim, K.C., Yang, X.H., Gu, J., Zhang, X.K., and Huang, S. (2004). A histone methyltransferase is required for maximal response to female sex hormones. Mol. Cell. Biol. *24*, 7032–7042.

Chan, C., Li, L., McCall, C.E., and Yoza, B.K. (2005). Endotoxin tolerance disrupts chromatin remodeling and NF-kappaB transactivation at the IL-1beta promoter. J. Immunol. *175*, 461–468.

Chen, J.D., and Evans, R.M. (1995). A transcriptional co-repressor that interacts with nuclear hormone receptors. Nature *377*, 454–457.

Culig, Z., Hobisch, A., Cronauer, M.V., Radmayr, C., Trapman, J., Hittmair, A., Bartsch, G., and Klocker, H. (1994). Androgen receptor activation in prostatic tumor cell lines by insulin-like growth factor-I, keratinocyte growth factor, and epidermal growth factor. Cancer Res. *54*, 5474–5478.

Dou, Y., Milne, T.A., Tackett, A.J., Smith, E.R., Fukuda, A., Wysocka, J., Allis, C.D., Chait, B.T., Hess, J.L., and Roeder, R.G. (2005). Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF. Cell *121*, 873–885.

Du, Y., Carling, T., Fang, W., Piao, Z., Sheu, J.C., and Huang, S. (2001). Hypermethylation in human cancers of the RIZ1 tumor suppressor gene, a member of a histone/protein methyltransferase superfamily. Cancer Res. *61*, 8094–8099.

Fischle, W., Wang, Y., and Allis, C.D. (2003). Binary switches and modification cassettes in histone biology and beyond. Nature *425*, 475–479.

Flanagan, J.F., Mi, L.Z., Chruszcz, M., Cymborowski, M., Clines, K.L., Kim, Y., Minor, W., Rastinejad, F., and Khorasanizadeh, S. (2005). Double chromodomains cooperate to recognize the methylated histone H3 tail. Nature *438*, 1181–1185.

Forneris, F., Binda, C., Vanoni, M.A., Battaglioli, E., and Mattevi, A. (2005). Human histone demethylase LSD1 reads the histone code. J. Biol. Chem. *280*, 41360–41365.

Forneris, F., Binda, C., Dall'Aglio, A., Fraaije, M.W., Battaglioli, E., and Mattevi, A. (2006). A highly specific mechanism of histone H3–K4 recognition by histone demethylase LSD1. J Biol Chem. *46*, 35289–35295.

Gazzerro, P., Abbondanza, C., D'Arcangelo, A., Rossi, M., Medici, N., Moncharmont, B., and Puca, G.A. (2006). Modulation of RIZ gene expression is associated to estradiol control of MCF-7 breast cancer cell proliferation. Exp. Cell Res. *312*, 340–349.

Guenther, M.G., Jenner, R.G., Chevalier, B., Nakamura, T., Croce, C.M., Canaani, E., and Young, R.A. (2005). Global and Hox-specific roles for the MLL1 methyltransferase. Proc. Natl. Acad. Sci. USA *102*, 8603–8608.

Hakimi, M.A., Bochar, D.A., Chenoweth, J., Lane, W.S., Mandel, G., and Shiekhattar, R. (2002). A core-BRAF35 complex containing histone deacetylase mediates repression of neuronal-specific genes. Proc. Natl. Acad. Sci. USA *99*, 7420–7425.

Hakimi, M.A., Dong, Y., Lane, W.S., Speicher, D.W., and Shiekhattar, R. (2003). A candidate X-linked mental retardation gene is a component of a new family of histone deacetylase-containing complexes. J. Biol. Chem. *278*, 7234–7239.

He, L., Yu, J.X., Liu, L., Buyse, I.M., Wang, M.S., Yang, Q.C., Nakagawara, A., Brodeur, G.M., Shi, Y.E., and Huang, S. (1998). RIZ1, but not the alternative RIZ2 product of the same gene, is underexpressed in breast cancer, and forced RIZ1 expression causes G2-M cell cycle arrest and/or apoptosis. Cancer Res. *58*, 4238–4244.

Heinzel, T., Lavinsky, R.M., Mullen, T.M., Soderstrom, M., Laherty, C.D., Torchia, J., Yang, W.M., Brard, G., Ngo, S.D., Davie, J.R., et al. (1997). A complex containing N-CoR, mSin3 and histone deacetylase mediates transcriptional repression. Nature *387*, 43–48.

Horlein, A.J., Naar, A.M., Heinzel, T., Torchia, J., Gloss, B., Kurokawa, R., Ryan, A., Kamei, Y., Soderstrom, M., Glass, C.K., et al. (1995). Ligand-independent repression by the thyroid hormone receptor mediated by a nuclear receptor co-repressor. Nature *377*, 397–404.

Humphrey, G.W., Wang, Y., Russanova, V.R., Hirai, T., Qin, J., Nakatani, Y., and Howard, B.H. (2001). Stable histone deacetylase complexes distinguished by the presence of SANT domain proteins CoREST/kiaa0071 and Mta-L1. J. Biol. Chem. *276*, 6817–6824.

Ishitani, T., Kishida, S., Hyodo-Miura, J., Ueno, N., Yasuda, J., Waterman, M., Shibuya, H., Moon, R.T., Ninomiya-Tsuji, J., and Matsumoto, K. (2003). The TAK1-NLK mitogen-activated protein kinase cascade functions in the Wnt-5a/Ca(2+) pathway to antagonize Wnt/beta-catenin signaling. Mol. Cell. Biol. *23*, 131–139.

Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. Science *293*, 1074–1080.

Kamei, Y., Xu, L., Heinzel, T., Torchia, J., Kurokawa, R., Gloss, B., Lin, S.C., Heyman, R.A., Rose, D.W., Glass, C.K., and Rosenfeld, M.G. (1996). A CBP integrator complex mediates transcriptional activation and AP-1 inhibition by nuclear receptors. Cell *85*, 403–414.

Kim, J., Jia, L., Tilley, W.D., and Coetzee, G.A. (2003). Dynamic methylation of histone H3 at lysine 4 in transcriptional regulation by the androgen receptor. Nucleic Acids Res. *31*, 6741–6747.

Kim, J., Jia, L., Stallcup, M.R., and Coetzee, G.A. (2005). The role of protein kinase A and cAMP response element-binding protein in androgen receptor-mediated transcription at the prostate-specific antigen locus. J. Mol. Endocrinol. *34*, 107–118.

Kouzarides, T. (2002). Histone methylation in transcriptional control. Curr. Opin. Genet. Dev. *12*, 198–209.

Kurahashi, T., Nomura, T., Kanei-Ishii, C., Shinkai, Y., and Ishii, S. (2005). The Wnt-NLK signaling pathway inhibits A-Myb activity by inhibiting the association with coactivator CBP and methylating histone H3. Mol. Biol. Cell *16*, 4705–4713.

Kwon, Y.-S., Garcia-Bassets, I., Hutt, K.R., Cheng, C.S., Jin, M., Liu, D., Benner, C., Wang, D., Ye, Z., Bibikova, M., Fan, J.-B., Duan, L., Glass, C.K., Rosenfeld, M.G., and Fu, X.-D. (2007). Sensitive ChIP-DSL technology reveals an extensive ERα binding program on human gene promoters. Proc. Natl. Acad. Sci. USA, in press. 10.1073/PNAS.0700715104.

Lachner, M., and Jenuwein, T. (2002). The many faces of histone lysine methylation. Curr. Opin. Cell Biol. *14*, 286–298.

Lee, M.G., Wynder, C., Cooch, N., and Shiekhattar, R. (2005). An essential role for CoREST in nucleosomal histone 3 lysine 4 demethylation. Nature *437*, 432–435.

Li, J., Lin, Q., Yoon, H.G., Huang, Z.Q., Strahl, B.D., Allis, C.D., and Wong, J. (2002). Involvement of histone methylation and phosphorylation in regulation of transcription by thyroid hormone receptor. Mol. Cell. Biol. *22*, 5688–5697.

Lunyak, V.V., Burgess, R., Prefontaine, G.G., Nelson, C., Sze, S.H., Chenoweth, J., Schwartz, P., Pevzner, P.A., Glass, C., Mandel, G., and Rosenfeld, M.G. (2002). Corepressor-dependent silencing of chromosomal regions encoding neuronal genes. Science *298*, 1747–1752.

Ma, H., Baumann, C.T., Li, H., Strahl, B.D., Rice, R., Jelinek, M.A., Aswad, D.W., Allis, C.D., Hager, G.L., and Stallcup, M.R. (2001). Hormone-dependent, CARM1-directed, arginine-specific methylation of histone H3 on a steroid-regulated promoter. Curr. Biol. *11*, 1981–1985.

Margueron, R., Trojer, P., and Reinberg, D. (2005). The key to development: interpreting the histone code? Curr. Opin. Genet. Dev. *15*, 163–176.

Martin, C., and Zhang, Y. (2005). The diverse functions of histone lysine methylation. Nat. Rev. Mol. Cell Biol. *6*, 838–849.

McKenna, N.J., and O'Malley, B.W. (2002). Combinatorial control of gene expression by nuclear receptors and coregulators. Cell *108*, 465–474.

McKinsey, T.A., Zhang, C.L., and Olson, E.N. (2001). Control of muscle development by dueling HATs and HDACs. Curr. Opin. Genet. Dev. *11*, 497–504.

Metivier, R., Penot, G., Carmouche, R.P., Hubner, M.R., Reid, G., Denger, S., Manu, D., Brand, H., Kos, M., Benes, V., and Gannon, F. (2004). Transcriptional complexes engaged by apo-estrogen receptor-alpha isoforms have divergent outcomes. EMBO J. *23*, 3653–3666.

Metzger, E., Wissmann, M., Yin, N., Muller, J.M., Schneider, R., Peters, A.H., Gunther, T., Buettner, R., and Schule, R. (2005). LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. Nature *437*, 436–439.

Nakamura, T., Mori, T., Tada, S., Krajewski, W., Rozovskaia, T., Wassell, R., Dubois, G., Mazo, A., Croce, C.M., and Canaani, E. (2002). ALL-1 is a histone methyltransferase that assembles a supercomplex of proteins involved in transcriptional regulation. Mol. Cell *10*, 1119–1128.

Narlikar, G.J., Fan, H.Y., and Kingston, R.E. (2002). Cooperation between complexes that regulate chromatin structure and transcription. Cell *108*, 475–487.

Nazareth, L.V., and Weigel, N.L. (1996). Activation of the human androgen receptor through a protein kinase A signaling pathway. J. Biol. Chem. *271*, 19900–19907.

Nielsen, S.J., Schneider, R., Bauer, U.M., Bannister, A.J., Morrison, A., O'Carroll, D., Firestein, R., Cleary, M., Jenuwein, T., Herrera, R.E., and

Kouzarides, T. (2001). Rb targets histone H3 methylation and HP1 to promoters. Nature *412*, 561–565.

Ogawa, H., Ishiguro, K., Gaubatz, S., Livingston, D.M., and Nakatani, Y. (2002). A complex with chromatin modifiers that occupies E2F- and Myc-responsive genes in G0 cells. Science *296*, 1132–1136.

Ogawa, S., Lozach, J., Jepsen, K., Sawka-Verheller, D., Perissi, V., Sasik, R., Rose, D.W., Johnson, R.S., Rosenfeld, M.G., and Glass, C.K. (2004). A nuclear receptor corepressor transcriptional checkpoint controlling activator protein 1-dependent gene networks required for macrophage activation. Proc. Natl. Acad. Sci. USA *101*, 14461–14466.

Perissi, V., Aggarwal, A., Glass, C.K., Rose, D.W., and Rosenfeld, M.G. (2004). A corepressor/coactivator exchange complex required for transcriptional activation by nuclear receptors and other regulated transcription factors. Cell *116*, 511–526.

Peterson, C.L., and Laniel, M.A. (2004). Histones and histone modifications. Curr. Biol. *14*, R546–R551.

Privalsky, M.L. (2004). The role of corepressors in transcriptional regulation by nuclear hormone receptors. Annu. Rev. Physiol. *66*, 315–360.

Rogatsky, I., Trowbridge, J.M., and Garabedian, M.J. (1999). Potentiation of human estrogen receptor alpha transcriptional activation through phosphorylation of serines 104 and 106 by the cyclin A-CDK2 complex. J. Biol. Chem. *274*, 22296–22302.

Rosenfeld, M.G., Lunyak, V.V., and Glass, C.K. (2006). Sensors and signals: a coactivator/corepressor/epigenetic code for integrating signal-dependent programs of transcriptional response. Genes Dev. *20*, 1405–1428.

Saccani, S., and Natoli, G. (2002). Dynamic changes in histone H3 Lys 9 methylation occurring at tightly regulated inducible inflammatory genes. Genes Dev. *16*, 2219–2224.

Schultz, D.C., Ayyanathan, K., Negorev, D., Maul, G.G., and Rauscher, F.J., 3rd. (2002). SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. Genes Dev. *16*, 919–932.

Shi, Y., Sawada, J., Sui, G., Affar, el B., Whetstine, J.R., Lan, F., Ogawa, H., Luke, M.P., Nakatani, Y., and Shi, Y. (2003). Coordinated histone modifications mediated by a CtBP co-repressor complex. Nature *422*, 735–738.

Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J.R., Cole, P.A., Casero, R.A., and Shi, Y. (2004). Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. Cell *119*, 941–953.

Shi, Y.J., Matson, C., Lan, F., Iwase, S., Baba, T., and Shi, Y. (2005). Regulation of LSD1 histone demethylase activity by its associated factors. Mol. Cell *19*, 857–864.

Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. Nature *403*, 41–45.

Steele-Perkins, G., Fang, W., Yang, X.H., Van Gele, M., Carling, T., Gu, J., Buyse, I.M., Fletcher, J.A., Liu, J., Bronson, R., et al. (2001). Tumor formation and inactivation of RIZ1, an Rb-binding member of a nuclear protein-methyltransferase superfamily. Genes Dev. *15*, 2250–2262.

Tong, J.K., Hassig, C.A., Schnitzler, G.R., Kingston, R.E., and Schreiber, S.L. (1998). Chromatin deacetylation by an ATP-dependent nucleosome remodelling complex. Nature *395*, 917–921.

Trewick, S.C., McLaughlin, P.J., and Allshire, R.C. (2005). Methylation: lost in hydroxylation? EMBO Rep. *6*, 315–320.

Tsukada, Y., Fang, J., Erdjument-Bromage, H., Warren, M.E., Borchers, C.H., Tempst, P., and Zhang, Y. (2006). Histone demethylation by a family of JmjC domain-containing proteins. Nature *439*, 811–816.

Ueda, T., Mawji, N.R., Bruchovsky, N., and Sadar, M.D. (2002). Ligand-independent activation of the androgen receptor by IL-6 and the role of the coactivator SRC-1 in prostate cancer cells. J. Biol. Chem. *277*, 38087–38094.

Wang, H., An, W., Cao, R., Xia, L., Erdjument-Bromage, H., Chatton, B., Tempst, P., Roeder, R.G., and Zhang, Y. (2003). mAM facilitates conversion by ESET of dimethyl to trimethyl lysine 9 of histone H3 to cause transcriptional repression. Mol. Cell *12*, 475–487.

Whetstine, J.R., Nottke, A., Lan, F., Huarte, M., Smolikov, S., Chen, Z., Spooner, E., Li, E., Zhang, G., Colaiacovo, M., and Shi, Y. (2006). Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylase. Cell *125*, 467–481.

Wysocka, J., Swigut, T., Milne, T.A., Dou, Y., Zhang, X., Burlingame, A.L., Roeder, R.G., Brivanlou, A.H., and Allis, C.D. (2005). WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. Cell *121*, 859–872.

Yamane, K., Toumazou, C., Tsukada, Y., Erdjument-Bromage, H., Tempst, P., Wong, J., and Zhang, Y. (2006). JHDM2A, a JmjC-containing H3K9 demethylase, facilitates transcription activation by androgen receptor. Cell *125*, 483–495.

Yang, L., Xia, L., Wu, D.Y., Wang, H., Chansky, H.A., Schubach, W.H., Hickstein, D.D., and Zhang, Y. (2002). Molecular cloning of ESET, a novel histone H3-specific methyltransferase that interacts with ERG transcription factor. Oncogene *21*, 148–152.

Yang, X.J., and Seto, E. (2003). Collaborative spirit of histone deacetylases in regulating chromatin structure and gene expression. Curr. Opin. Genet. Dev. *13*, 143–153.

You, A., Tong, J.K., Grozinger, C.M., and Schreiber, S.L. (2001). CoREST is an integral component of the CoREST- human histone deacetylase complex. Proc. Natl. Acad. Sci. USA *98*, 1454–1458.

Zhang, C.L., McKinsey, T.A., and Olson, E.N. (2002). Association of class II histone deacetylases with heterochromatin protein 1: potential role for histone methylation in control of muscle differentiation. Mol. Cell. Biol. *22*, 7302–7312.

Zhu, P., Baek, S.H., Bourk, E.M., Ohgi, K.A., Garcia-Bassets, I., Sanjo, H., Akira, S., Kotol, P.F., Glass, C.K., Rosenfeld, M.G., and Rose, D.W. (2006). Macrophage/cancer cell interactions mediate hormone resistance by a nuclear receptor derepression pathway. Cell *124*, 615–629.

Zwijsen, R.M., Wientjens, E., Klompmaker, R., van der Sman, J., Bernards, R., and Michalides, R.J. (1997). CDK-independent activation of estrogen receptor by cyclin D1. Cell *88*, 405–415.

Sensitive ChIP-DSL technology reveals an extensive estrogen receptor α-binding program on human gene promoters

Young-Soo Kwon*, Ivan Garcia-Bassets†, Kasey R. Hutt†‡, Christine S. Cheng†‡, Mingjie Jin§, Dongyan Liu§, Chris Benner*‡, Dong Wang*, Zhen Ye*, Marina Bibikova¶, Jian-Bing Fan¶, Lingxun Duan§, Christopher K. Glass*, Michael G. Rosenfeld†‖, and Xiang-Dong Fu*‖

*Department of Cellular and Molecular Medicine, University of California at San Diego School of Medicine, La Jolla, CA 92093-0651; †Department of Medicine, Howard Hughes Medical Institute and University of California at San Diego School of Medicine, La Jolla, CA 92093; ‡Bioinformatics Graduate Program, University of California at San Diego, La Jolla, CA 92093; §Aviva Systems Biology Corporation, San Diego, CA 92121; and ¶Illumina Inc., San Diego, CA 92121

Contributed by Michael G. Rosenfeld, January 26, 2007 (sent for review January 22, 2007)

ChIP coupled with microarray provides a powerful tool to determine *in vivo* binding profiling of transcription factors to deduce regulatory circuitries in mammalian cells. Aiming at improving the specificity and sensitivity of such analysis, we developed a new technology called ChIP-DSL using the DNA selection and ligation (DSL) strategy, permitting robust analysis with much reduced materials compared with standard procedures. We profiled general and sequence-specific DNA binding transcription factors using a full human genome promoter array based on the ChIP-DSL technology, revealing an unprecedented number of the estrogen receptor (ERα) target genes in MCF-7 cells. Coupled with gene expression profiling, we found that only a fraction of these direct ERα target genes were highly responsive to estrogen and that the expression of those ERα-bound, estrogen-inducible genes was associated with breast cancer progression in humans. This study demonstrates the power of the ChIP-DSL technology in revealing regulatory gene expression programs that have been previously invisible in the human genome.

breast cancer | genome-wide | promoter array

The elucidation of genomes for humans and other model organisms has made it possible to conduct analysis of gene expression and regulation at the genome scale. Gene expression is generally accompanied by chromatin remodeling activities and histone modifications. An important conceptual advance has been the "histone code" hypothesis, which suggests that histone modifications reflect a sequential action of enzymes associated with the transcriptional machinery such that one prior activity may influence the next during regulated gene expression (1, 2). Histone acetylation results in charge neutralization of modified lysines, which is generally associated with gene activation (3). In contrast, histone methylation on different residues appears to provide binding sites for specific transcription regulators, thereby positively or negatively affecting gene expression (4). Although histone methylation may modulate gene expression in a gene-specific and context-dependent manner, certain site-specific modifications appear to be generally applicable to most genes. The epigenetic markers thus provide a roadmap to identify and characterize functional DNA elements in the genome.

The nuclear receptor (NR) superfamily of transcriptional regulators plays a central role in many developmental and disease processes, and the system has been extensively studied as a model to learn the mechanism for spatial and temporal control of gene expression (5). Individual NRs have consensus binding sites in promoters and enhancers, which have been characterized in detail, but only in a limited number of NR-regulated genes. In the case of the *pS2* gene (also known as *TFF1*), for example, binding by estrogen receptor α (ERα) initiates sequential recruitment of a large number of transcription factors onto the promoter to start transcription (6). However, despite extensive mechanistic insights in transcriptional initiation in this and other well studied cases, little is known about how many genes are direct targets for an NR. Genome-wide ChIP coupled with microarray, known as ChIP-on-chip, offers a solution to this problem by determining promoters bound directly by transcription factors (7–10). Surprisingly, however, recent promoter and tiling array analyses suggest that ERα binds relatively rarely to gene promoters compared with intergenic regions, suggesting a critical role of long-distance enhancers in regulated gene expression in mammalian cells (11–13).

Here we describe an approach to detecting *in vivo* DNA–protein interactions by coupling ChIP with a DNA selection and ligation (DSL) strategy, permitting analysis of many fewer cells than required by the conventional ChIP-on-chip method. We constructed a full genome promoter array based on this ChIP-DSL platform, and our analysis revealed that ERα bound to >3% of human genes in promoter-proximal regions in MCF-7 cells, reinforcing the importance of direct binding events in the promoter-proximal regions during regulated gene expression. Results from built-in tiling arrays allowed direct visualization of binding events even without statistical filtering of raw data, and a comprehensive histone modification profile extended the current histone code hypothesis. These results demonstrate the versatility and accuracy of the ChIP-DSL technology in a genome-wide search for direct target genes by specific transcription factors and in comprehensive analysis of regulatory programs within specific genomic loci. Furthermore, comparison between profiles of ERα binding and 17β-estradiol (E_2)-induced gene expression in MCF-7 cells revealed a subset of genes whose expression tracks breast cancer progression in humans, which not only suggests the prognostic value of these genes as biomarkers for breast cancer but also illustrates a general strategy for dissecting molecular pathways in cancer.

Author contributions: Y.-S.K. and I.G.-B. contributed equally to this work; Y.-S.K., I.G.-B., K.R.H., C.K.G., M.G.R., and X.-D.F. designed research; Y.-S.K., I.G.-B., K.R.H., D.W., M.B., and J.-B.F. performed research; M.J., D.L., Z.Y., and L.D. contributed new reagents/analytic tools; K.R.H., C.S.C., and C.B. analyzed data; and I.G.-B., K.R.H., M.G.R., and X.-D.F. wrote the paper.

M.J., D.L., and L.D. are employees of Aviva Systems Biology Corporation. X.-D.F. is scientific founder of Aviva Systems Biology Corporation and is a member of its board of directors. M.B. and J.-B.F. are employees of Illumina Inc. Both companies market array products used in this paper.

Freely available online through the PNAS open access option.

Abbreviations: DSL, DNA selection and ligation; E_2, 17β-estradiol; Pol, polymerase; qPCR, quantitative PCR; NR, nuclear receptor.

Data deposition: The genomic coordinates for the annotated human gene promoters and the array data have been deposited in the ArrayExpress database, www.ebi.ac.uk/aerep (accession nos. E-MEXP-984 and E-TABM-231).

‖To whom correspondence may be addressed. E-mail: mrosenfeld@ucsd.edu or xdfu@ucsd.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0700715104/DC1.

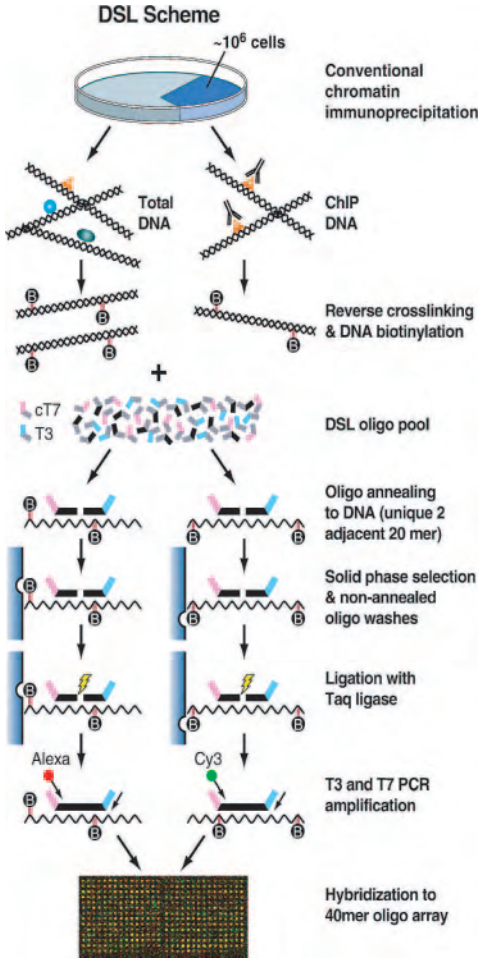

Fig. 1. The ChIP-DSL scheme. A key feature of the technology is oligonucleotide ligation templated by chromatin immunoprecipitated (ChIP) DNA followed by DSL. This permits high-throughput analysis of target genes with much improved specificity and sensitivity.

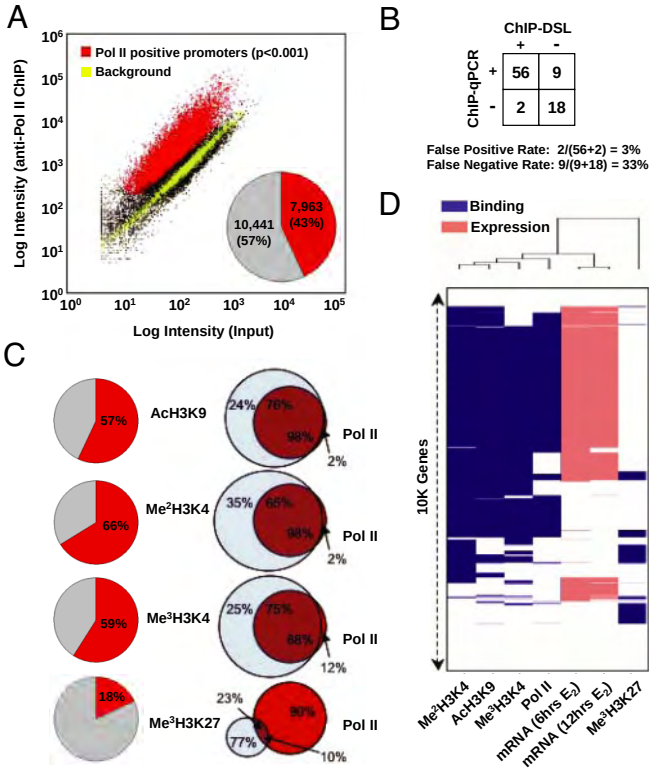

Fig. 2. Global analysis of promoter occupancy by ChIP-DSL. (*A*) Global analysis of Pol II-bound promoters in E_2-stimulated MCF-7 cells. A set of tiled genomic loci (yellow) served as internal negative controls because most genomic sequences are not expected to interact with general and sequence-specific transcription factors. Pol II-positive (red) and -negative (black) promoters were identified based on the single-array error model at $P < 0.001$, and the percentages of Pol II-positive and -negative promoters are shown in *Inset*. (*B*) ChIP/qPCR validation of the ChIP-DSL results. (*C*) Promoter profiling of modified histones in E_2-treated MCF-7 cells. (*Left*) Percentages of positive promoters. (*Right*) Overlap of positive promoters with Pol II. The overlap between Pol II binding and individual histone modification events is shown in individual Venn diagrams. (*D*) Correlation of gene expression with promoter occupancy by Pol II and histone modification markers. Gene expression profiling in E_2-induced MCF-7 cells was carried out on Illumina gene expression arrays. Approximately 10,000 genes common to both promoter and expression profiling arrays and reliably scored in all measurements were used to construct the binary map by unsupervised hierarchical clustering analysis.

Results and Discussion

Design and Development of the ChIP-DSL Technology. Aiming to detect functional DNA elements with high sensitivity and specificity, we devised a multiplex assay by coupling ChIP with a DSL approach (Fig. 1). A signature 40-nt sequence is first computationally identified in a genomic segment 0.5–1 kb in length. For promoter profiling each such probe corresponds to a proximal promoter region from +200 nt to −800 nt relative to the transcription start, which contains ≈95% of known binding sites for transcription factors in humans (8). To construct a tiling array, each probe is used to represent an ≈0.5-kb nonrepetitive genomic block in a path to be tiled. This probe density takes into account the number of probes required for maximal coverage of genomic sequences and the sufficiency in detecting immunoprecipitated DNA, which is generally sheared to an average length of 0.5–1 kb. Amine-modified 40-mers are spotted onto solid support to form an array.

Corresponding to each 40-mer, a pair of assay oligonucleotides are synthesized, each consisting of the two 20-mer halves in the 40-mer and flanked by a universal primer landing site. Multiple oligonucleotide pairs are mixed to form a pool. The assay begins with standard ChIP, and the isolated DNA is randomly biotinylated followed by annealing to the oligonucleotide pool. Annealed oligonucleotides are selected on streptavidin-conjugated magnetic beads, and unannealed oligonucleotides are washed away. This selection strategy allows the use of an excessive amount of oligonucleotides to achieve maximal annealing that follows the pseudo-first-order kinetics and prevents interference of later steps by excess free oligonucleotides in solution. All selected oligonucleotides are immobilized, and those paired by specific target DNA are ligated, thereby converting only correctly targeted oligonucleotides to full amplicons for PCR amplification. One of the PCR primers is end-labeled with a fluorescent dye so that the PCR products can be directly hybridized to the 40-mer array.

This technology is distinct from the conventional ChIP-on-chip assay in several key aspects. First, chromatin immunoprecipitated DNA was used to template oligonucleotide ligation, instead of being directly amplified for hybridization. This step can tolerate incomplete decross-linking because cross-linking adducts should have less effect on oligonucleotide hybridization than signal amplification by a polymerase. Second, we targeted only unique signature sequences in the human genome, thereby avoiding potential interference of repetitive and related sequences during hybridization. Third, the sensitivity is significantly elevated by PCR amplification of ligated oligonucleotides in an unbiased fashion because all amplicons contain the same pair of specific primer

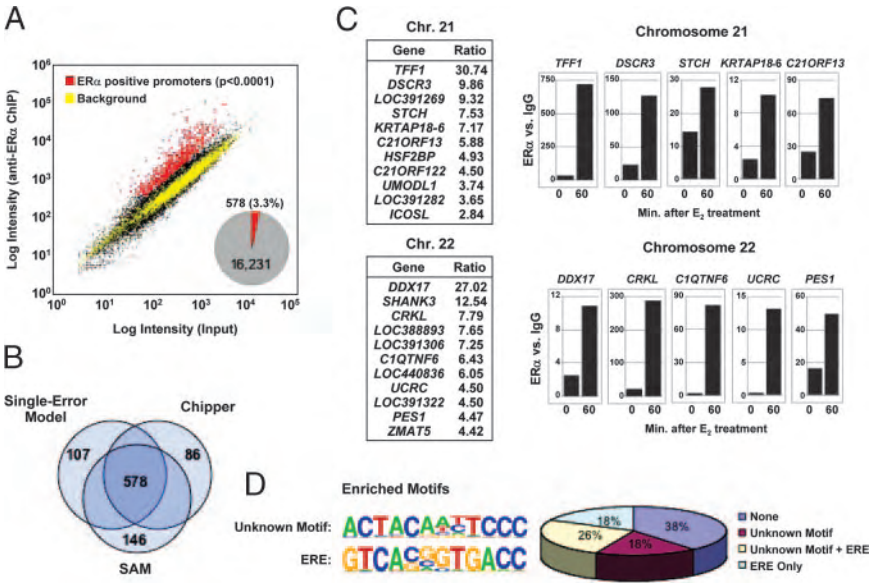

Fig. 3. Promoter profiling of ERα in E_2-induced MCF-7 cells. (*A* and *B*) ERα-bound promoters were identified (red) at $P < 0.0001$. The percentage of ERα-bound promoters scored positively by all three analytical methods is shown in *A Inset*, and additional promoters scored positively by one or two methods are indicated in the peripheries of the Venn diagram in *B*. (*C*) Listed are the newly identified ERα-positive promoters on chromosomes 21 and 22. Ratios were deduced from array measurements. Selected promoters were validated by ChIP/qPCR (*Right*). (*D*) Motif analysis of anti-ERα-enriched promoters. The first motif appears common to gene promoters in general, but the protein(s) recognizing this motif is unknown. When this motif is masked, the most enriched motif corresponds to full- or half-consensus estrogen responsive element. Allowing one base mismatch, the percentage of promoters containing a full- or half-consensus estrogen responsive element among total ERα-bound promoters was calculated and shown in *Right*.

landing sites and are uniform in length as previously documented (14).

We progressively enlarged the multiplicity of the assay to eventually cover most annotated gene promoters in the human genome. Titration experiments indicate that the ChIP-DSL technology could routinely operate with cells from one-third of a single 100-mm culture dish, which corresponds to 1–5 × 10^6 cells, depending on the cell type under investigation. Despite the fact that each promoter is targeted by one oligonucleotide pair, high-quality data generated as reported in this and other studies (15) demonstrate the reproducibility and robustness of the ChIP-DSL technology.

We initially assessed promoters potentially active in transcription based on their association with RNA polymerase (Pol) II in E_2-treated MCF-7 cells (Fig. 2*A*). Anti-Pol II-enriched (Pol II^+) promoters (red) were clearly segregated from background marked by built-in tiling array controls (yellow), finding that 43% of total promoters were Pol II^+ at a standard P value < 0.001. Quantitative ChIP/quantitative PCR (qPCR) analysis of randomly selected promoters suggests a false positive rate of ≈3% and a false negative rate of ≈33% [Fig. 2*B* and supporting information (SI) Fig. 6]. A similar false positive rate was observed by using an irrelevant IgG (SI Fig. 7*A*). The false negative rate is quite similar to that reported in published ChIP-chip studies (16). Pol II^+ promoters were also marked by AcH3K9 (98%), Me^2H3K4 (98%), and Me^3H3K4 (88%), although a significant fraction of promoters were associated only with these gene activation marks, but not with Pol II (Fig. 2*C*). In contrast, the "repressive" histone mark Me^3H3K27 was detected in only a small fraction (10%) of Pol II^+ promoters. Indeed, this repressive histone mark has been shown to associate with some active genes (17). An RNA profiling experiment in the same E_2-stimulated MCF-7 cells showed that most Pol II^+ promoters were actively transcribing (Fig. 2*D*). Collectively, these robust and highly consistent data testify to the utility and sensitivity of the ChIP-DSL technology.

Identification of ERα-Occupied Gene Promoters in the Human Genome. We next applied the ChIP-DSL technology to identify target genes for sequence-specific DNA-binding transcription factors. ERα plays an important role in human reproduction and breast cancer. Recent promoter profiling analysis using 1-kb promoter regions detected 153 ERα-bound promoters (13). Further, tiling analysis of ERα binding suggests that ERα binds prevalently to intergenic regions in the human genome, suggesting a new paradigm that estrogen-regulated gene expression may be largely driven by long-distance enhancers (11, 12). Complimentary to these recent genomic analyses, we scored ≈1,300 anti-ERα-enriched ($ERα^+$) promoters in E_2-stimulated MCF-7 cells based on the single-array error model (9) at the standard cutoff of $P < 0.001$, and ≈700 at a more stringent cutoff of $P < 0.0001$ (Fig. 3*A*). A significant number of $ERα^+$ promoters were also identified in vehicle-treated MCF-7 cells, suggesting a class of hormone-independent recruitment events (SI Fig. 7*B*). To identify $ERα^+$ promoters with high statistical confidence, we analyzed the data from multiple biological repeats using three statistical methods that are based on distinct mathematical principles, revealing an overlapping set of 578 highest confidence $ERα^+$ promoters, which represents 3.3% of all reliably scored promoters (Fig. 3*B* and SI Data Set 1).

ChIP/qPCR analysis confirmed all $ERα^+$ promoters examined, including those residing in chromosomes 21 and 22 (Fig. 3*C*) and 20 additional promoters in other chromosomes (data not shown), indicating a negligible false positive rate for anti-$ERα^+$-enriched promoters supported by stringent statistical tests. Estimation of the false negative rate proved to be challenging when the majority of the probes are in the "negative" population (8). We used ChIP/qPCR-confirmed promoters recently reported (13) to objectively estimate our false negative rate. Among 27 validated promoters common between the two array platforms, 20 were scored positive in our array at $P < 0.0001$ and 24 at $P < 0.001$, indicating that our false negative rate is ≈26% and 11% at the two P value cutoffs, respectively, which is probably an overestimate, because three promoters (*CYP4F3*, *PROP1*, and *ABCG2*) not detected were enriched only <2-fold in previous ChIP/qPCR experiments (13). Together, these results demonstrate the accuracy of the ChIP-DSL data and conservatively identify ≈4-fold as many ERα target promoters as were detected in previous genome-wide location analysis, suggesting that the promoter array based on the ChIP-DSL technology is a useful resource for the general research community.

We next conducted motif analysis using a newly refined algorithm, which compares ChIP-enriched promoters against normalized nucleotide frequencies in all promoters (C.B. and C.K.G., unpublished data), revealing a highly enriched, but uncharacterized, motif associated with ERα-bound promoters (Fig. 3*D*). When this motif was masked, the next most enriched motifs were the classic ERα-binding consensus sequences (18), which are present in 44% of the total $ERα^+$ promoters (Fig. 3*D*). Interestingly, whereas the algorithm confirmed the presence of FoxA1 recognition motifs surrounding a fraction of intergenic ERα binding sites (11), it did not detect extensive association of the FoxA1 binding site with the

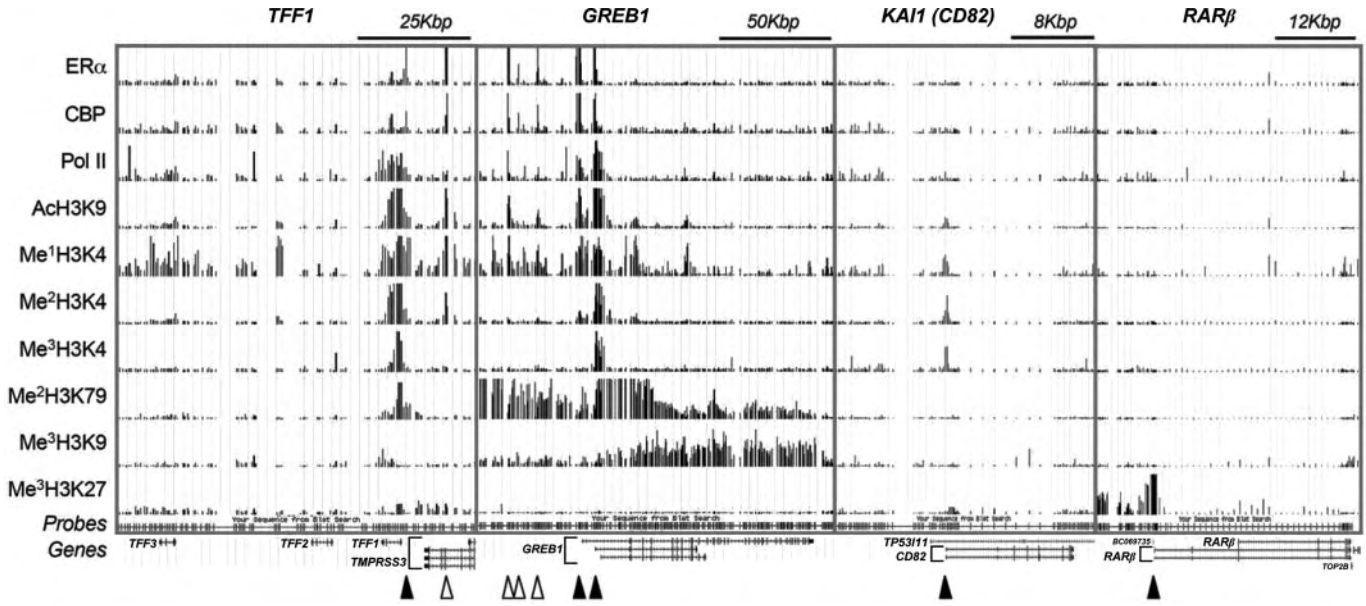

Fig. 4. Locus-specific tiling array analysis of ERα binding and histone modifications in E_2-induced MCF-7 cells. Individual genes and scales are shown at the top, and probe positions and gene structure are indicated at the bottom. Individual transcription factors and chromatin remodeling markers profiled are indicated on the left. Transcription starts and known or putative enhancers are designated by filled and open arrowheads, respectively, at the bottom.

ERα$^+$ promoters identified by ChIP-DSL. In light of the finding that FoxA1 is critical for ERα binding to several target genes examined (11, 13), it will be interesting to determine in future studies whether FoxA1 is selectively or universally required for ERα targeting.

Locus-Specific Tiling Array Analysis of ERα Binding and Histone Modifications. To facilitate data analysis of promoter arrays, we built in a number of tiled loci to serve as internal negative controls because not all genomic regions are expected to be occupied by general and sequence-specific DNA binding transcription factors. The data in turn illustrate the usefulness of the ChIP-DSL technology in revealing specific molecular recognition events that constitute the regulatory programs in individual genomic loci. As illustrated in Fig. 4, we found that ERα bound to the promoter (filled arrow) and a putative enhancer (open arrow) of the *TFF1* gene, as previously reported (11). The transcriptional coactivator CBP similarly interacted with both promoter and enhancer, whereas Pol II covered the body of this relatively small gene. In the case of the *GREB1* gene, we observed a similar pattern with ERα, CBP, and Pol II present on two of the three promoters that were previously characterized (19). Interestingly, we found that all three factors interacted with three distinct loci upstream of the *GREB1* promoters, suggesting that these sites may function as enhancers. These observations are consistent with a large body of literature that gene promoters and enhancers are recognized by sequence-specific DNA-binding transcription factors, which in turn recruit transcription coactivators.

Acetylated histone (AcH3K9) was observed in both promoters and enhancers in *TFF1* and *GREB1* as expected. Histone 3 lysine-4 methylation is generally associated with active genes, but the profile of individual modifications is significantly distinct: Me^{1}H3K4 seems to associate broadly with active genes, but, in contrast to the situation in yeast, this modification is not preferentially linked to the 3′ end of active genes (2, 20, 21). Me^{2}H3K4 marks both promoters and enhancers with a clear preference for promoters over enhancers. Again, in contrast to events in yeast, we did not detect substantial Me^{2}H3K4 in the transcribed regions of *GREB1* and other tiled genes. Me^{3}H3K4 was found exclusively in promoters, which agrees with most mapping studies in yeast and mammalian cells (20, 22, 23).

Interestingly, AcH3K9 and methylated H3K4 marks were present in a number of gene promoters, including *KAI1* (Fig. 4), where no RNA transcripts were detected in MCF-7 cells. These observations suggest either that some histone modifications take place before the recruitment of the general transcriptional machinery or that these genes may be transcribing at an undetectable basal level. Although one cannot formally distinguish between these possibilities, the histone modification pattern is clearly different from "silent" genes such as *RARβ* (Fig. 4). The *RARβ* promoter was specifically marked by Me^{3}H3K27, which is generally associated with silent genes in heterochromatin (24). Thus, there is heterogeneity in marks of nonexpressed genes exemplified by the observation that the *KAI1* promoter is accessible to transcription factors, whereas the *RARβ* promoter is actively repressed in MCF-7 cells.

To further characterize histone modifications associated with gene repression, we mapped Me^{2}H3K79 and Me^{3}H3K9 (Fig. 4). Me^{2}H3K79 has been previously implicated in interactions with Sir proteins during gene silencing (25, 26), although a more recent study suggested a link of this modification to gene activation (27). We found that Me^{2}H3K79 was indeed associated with active genes, but in a distinct, gene-specific manner. In the case of *TFF1*, this modification took place in the transcribing region near the promoter, whereas in the case of *GREB1*, Me^{2}H3K79 was spread in the entire transcription unit, including both coding and promoter/enhancer regions. Me^{3}H3K9 has also been previously linked to gene repression by serving as the binding site for HP1 to facilitate the assembly of heterochromatin (25, 26, 28–30). Here we found that Me^{3}H3K9 decorated most 3′ transcribed regions of both *TFF1* and *GREB1*, consistent with a role of this specific histone modification in transcription elongation as recently suggested in yeast (31). These findings illustrate that it is still quite precipitous to generalize the significance of most histone modification events with respect to gene activation or repression, as the same histone modification may reflect or influence transcription positively or negatively in a highly gene-specific and locus-dependent manner, consistent with a combinatorial histone code.

Expression of Direct ERα Target Genes in Breast Cancer Cells and Tissues. While the majority of ERα-bound promoters were also marked by Pol II and epigenetic markers associated with gene

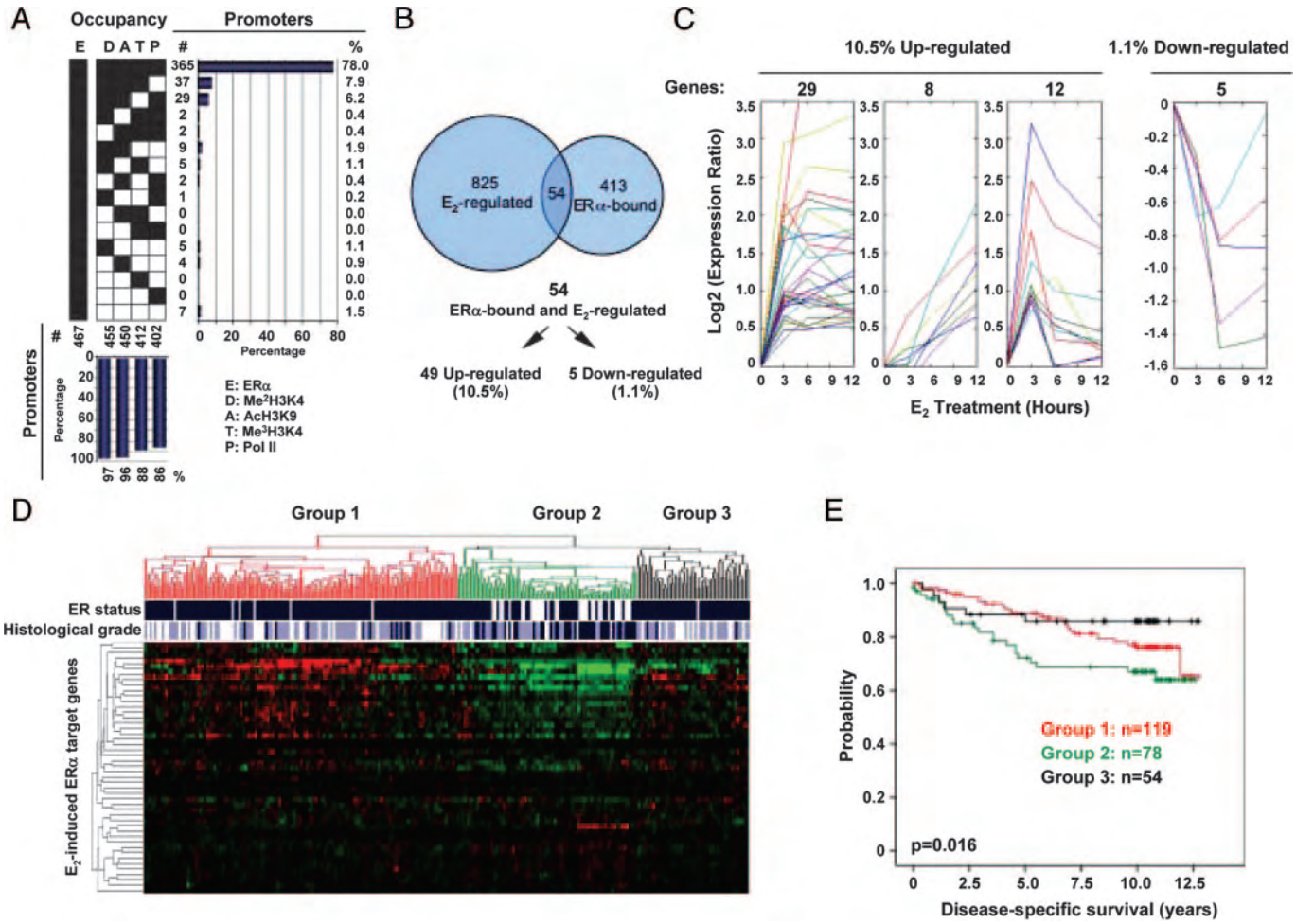

Fig. 5. E_2-induced gene expression and the biological relevance of direct ERα target genes. (*A*) Relationship between ERα binding and histone modifications. To directly compare ERα binding and E_2-induced gene expression, 467 of 578 ERα-bound promoters common between our promoter array and the Illumina gene expression array were analyzed. The majority of ERα-positive promoters was also marked by Pol II and modified histones associated with gene activation. (*B*) Venn diagram showing the overlap between ERα-bound promoters and E_2-induced genes. (*C*) Gene expression profiling in response to E_2 treatment. ERα-bound and E_2-regulated genes are grouped into four distinct classes. Among up-regulated genes, 29 were rapidly induced, and the level remained relatively constant afterward; eight were induced in a time-dependent manner; and 12 were induced followed by a rapid decay. E_2-induced, genes represent 10.5% of total ERα-bound genes in the promoter-proximal region. Only five ERα-bound genes were down-regulated by E_2, which represent 1.1% of total ERα-bound genes in the promoter-proximal region. (*D*) Segregation of ER expression and breast tumor grade (both indicated at the top by blue bars) based on ERα-bound and E_2-induced genes in MCF-7 cells. (*E*) Kaplan–Meier plots of patient survival in different groups segregated based on ERα-bound and E_2-inducted genes in MCF-7 cells. Statistical significance was determined by the χ^2 test.

activation (Fig. 5*A*), we evaluated the time course of regulated gene expression by RNA profiling and identified 879 genes that responded to E_2-induction in MCF-7 cells, which generally agrees with other published gene expression profiling studies (32–34). Strikingly, only 54 of these 879 E_2-affected genes were bound by ERα in the promoter-proximal region (Fig. 5*B*), indicating that the majority of E_2-induced genes might be indirectly affected or regulated by ER-responsive elements located away from the promoter-proximal region. Among these 879 E_2-regulated genes, 562 were up-regulated and 317 were down-regulated. Contrary to the expectation that a similar percentage of genes in these two categories would be targeted by ERα in the promoter-proximal region, we found that 49 (10.5%) of ERα-bound promoters were up-regulated by E_2, whereas only 5 (1.1%) were down-regulated (Fig. 5 *B* and *C*). These observations suggest that many genes in both up- and down-regulated categories might be indirectly affected, with more down-regulated genes influenced by indirect mechanisms than up-regulated ones (35).

Conversely, the fact that only 54 ERα-bound promoters responded to E_2 with rapid changes in mRNA levels suggests that most ERα-bound promoters may require additional cofactors for E_2-dependent gene expression, as has been previously documented (36–39). Consequently, we predicted that different sets of ERα-occupied promoters might respond to E_2 stimulation in different cell types. Indeed, we have found that a subset of ERα-bound, but E_2-insensitive, promoters in MCF-7 cells could be directly targeted by ERα and induced by E_2 in U2OS cells stably expressing ERα (data not shown). This observation suggests that, at least for some promoters, they represent bona fide estrogen target genes under different circumstances.

To further investigate the biological relevance of ERα-binding and estrogen-regulated gene expression, we asked how the newly identified 54 E_2-responsive ERα target genes might be differentially regulated in breast cancer tissues using a comprehensive set of gene expression profiling data from 251 breast cancer patients (40). We found a direct correlation between gene expression and tumor progression by unsupervised hierarchical clustering (Fig. 5*D*). Patients were clustered into three groups. About half of the genes were strongly suppressed in group 2, which displayed an ER-negative status and advanced tumor grade (Fig. 5*D*). Significantly, this

patient group exhibited a much reduced survival rate compared with the two other groups (Fig. 5*E*). These results illustrate a general strategy for disease etiology studies by combining gene expression profiling with location analysis of key transcriptional regulators altered in specific diseases.

Materials and Methods

Cell Culture and Antibodies. MCF-7 cells were cultured in MEM supplemented with 10% FBS. Before induction, cells were hormone-deprived for 4 days in phenol-free MEM plus charcoal-depleted FBS and then treated with 100 nM E_2 (Sigma–Aldrich, St. Louis, MO) for 1 h for ChIP or various periods of time for RNA profiling as indicated. Antibodies used for ChIP analyses were anti-RNAP (8WG16) (MMS-126R; Covance, Princeton, NJ), anti-ERα (HC-20 and H-184 combined; Santa Cruz Biotechnology, Santa Cruz, CA), anti-CBP (C-20 and A22 combined; Santa Cruz Biotechnology). All anti-modified histone antibodies are from Upstate Biotechnology (Lake Placid, NY), including anti-AcH3K9 (07-352), anti-Me^{1}H3K4 (07-436), anti-Me^{2}H3K4 (07-030), anti-Me^{3}H3K4 (07-473), anti-Me^{3}H3K9 (07-442), anti-Me^{3}H3K27 (07-449), and anti-Me^{2}H3K79 (07-366).

Array Fabrication and the ChIP-DSL Assay. Human promoters were annotated by aligning Refseq mRNAs against the human genome and extended by using existing ESTs. A sequence from +200 to −800 bp relative to each transcription start was used to determine the most unique 40-mer to represent that promoter. All 40-mer oligonucleotides were amino-derived during oligo synthesis and printed on the 3D-CodeLink slides according to the manufacturer's instructions (Amersham Biosciences). Corresponding to each 40-mer, a pair of assay oligonucleotides were synthesized, each containing a half of the 40-mer sequence, flanked by a universal primer binding site. The built-in tiling paths for internal controls were based on sequences from multiple human genes, and oligonucleotides probes were selected at the ≈0.5-kb interval across each gene unit. The genomic coordinates for the annotated human gene promoters and the array data have been submitted to ArrayExpress (www.ebi.ac.uk/aerep).

Cells were cross-linked by formaldehyde and subjected to standard ChIP as previously described (41). Cells in one 100-mm dish were used for each ChIP-DSL experiment. Both input (≈5% of total DNA) and antibody-enriched DNA were randomly biotinylated by using a kit (Vector Laboratories) according to the manufacturer's instructions. All T7-linked assay oligonucleotides were kinased and then mixed with all T3-linked oligonucleotides. For each reaction, we used 0.1 pmol per oligonucleotide in a pool suspended in 10 μl of TE buffer. The procedure for oligonucleotide annealing, solid phase selection, ligation, and PCR amplification was as described (42), except *Taq* ligase was used in place of T4 ligase to improve ligation specificity. Input DNA was labeled with Alexa Fluor 647 and chromatin immunoprecipitated DNA with Cy3. The PCR products were mixed, denatured, and hybridized to the 40-mer Hu20K array. Slides were scanned on the GenPix 4000B scanner (Axon Instruments). The Hu20K array and the associated assay kit with detailed instruction are commercially available from Aviva Systems Biology.

Data Analysis. The single-array error model was previously described (43, 44). The SAM analysis package (www-stat.stanford.edu/~tibs/SAM) was previously described (45). Chipper (http://llama.med.harvard.edu/cgi/Chipper/chip3.py?id=725676) was described (46). After conducting analysis with these methods, we first obtained genes at $P < 0.0001$ according to the single-error model and selected the same number of genes from the top list of the other two methods to identify genes that were scored significantly by all three methods. Clearly, genes identified by only one or two of the methods may still be highly significant.

We are indebted to Bing Ren for generous help during the course of the technology development. We thank colleagues in our laboratories, especially V. Lunyak for insightful discussion and advice and C. Nelson for cell culture. M.G.R. is an Investigator of the Howard Hughes Medical Institute. This work was supported by grants from the National Institutes of Health and the Vitamin Cases Consumer Settlement Fund (to M.G.R.), and by National Cancer Institute Grant CA114184 and National Human Genome Research Institute Grant HG003119 (to X.-D.F.).

1. Strahl BD, Allis CD (2000) *Nature* 403:41–45.
2. Fischle W, Wang Y, Allis CD (2003) *Curr Opin Cell Biol* 15:172–183.
3. Turner BM (2000) *BioEssays* 22:836–845.
4. Sims RJ, III, Nishioka K, Reinberg D (2003) *Trends Genet* 19:629–639.
5. Rosenfeld MG, Lunyak VV, Glass CK (2006) *Genes Dev* 20:1405–1428.
6. Metivier R, Penot G, Hubner MR, Reid G, Brand H, Kos M, Gannon F (2003) *Cell* 115:751–763.
7. Odom DT, Zizlsperger N, Gordon DB, Bell GW, Rinaldi NJ, Murray HL, Volkert TL, Schreiber J, Rolfe PA, Gifford DK, *et al.* (2004) *Science* 303:1378–1381.
8. Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS, Zucker JP, Guenther MG, Kumar RM, Murray HL, Jenner RG, *et al.* (2005) *Cell* 122:947–956.
9. Ren B, Robert F, Wyrick JJ, Aparicio O, Jennings EG, Simon I, Zeitlinger J, Schreiber J, Hannett N, Kanin E, *et al.* (2000) *Science* 290:2306–2309.
10. Harbison CT, Gordon DB, Lee TI, Rinaldi NJ, Macisaac KD, Danford TW, Hannett NM, Tagne JB, Reynolds DB, Yoo J, *et al.* (2004) *Nature* 431:99–104.
11. Carroll JS, Liu XS, Brodsky AS, Li W, Meyer CA, Szary AJ, Eeckhoute J, Shao W, Hestermann EV, Geistlinger TR, *et al.* (2005) *Cell* 122:33–43.
12. Carroll JS, Meyer CA, Song J, Li W, Geistlinger TR, Eeckhoute J, Brodsky AS, Keeton EK, Fertuck KC, Hall GF, *et al.* (2006) *Nat Genet* 38:1289–1297.
13. Laganiere J, Deblois G, Lefebvre C, Bataille AR, Robert F, Giguere V (2005) *Proc Natl Acad Sci USA* 102:11651–11656.
14. Fan JB, Yeakley JM, Bibikova M, Chudin E, Wickham E, Chen J, Doucet D, Rigault P, Zhang B, Shen R, *et al.* (2004) *Genome Res* 14:878–885.
15. Garcia-Bassets I, Kwon Y-S, Telese F, Perfontaine GG, Hutt KR, Cheng CS, Ju B-G, Ohgi KA, Wang J, Escoubet-Lozach L, *et al.* (2007) *Cell* 128:505–518.
16. Kim TH, Barrera LO, Zheng M, Qu C, Singer MA, Richmond TA, Wu Y, Green RD, Ren B (2005) *Nature* 436:876–880.
17. Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, *et al.* (2006) *Cell* 125:315–326.
18. Klinge CM (2001) *Nucleic Acids Res* 29:2905–2919.
19. Ghosh MG, Thompson DA, Weigel RJ (2000) *Cancer Res* 60:6367–6375.
20. Pokholok DK, Harbison CT, Levine S, Cole M, Hannett NM, Lee TI, Bell GW, Walker K, Rolfe PA, Herbolsheimer E, *et al.* (2005) *Cell* 122:517–527.
21. Schubeler D, Turner BM (2005) *Cell* 122:489–492.
22. Ng HH, Robert F, Young RA, Struhl K (2003) *Mol Cell* 11:709–719.
23. Bernstein BE, Kamal M, Lindblad-Toh K, Bekiranov S, Bailey DK, Huebert DJ, McMahon S, Karlsson EK, Kulbokas EJ, III, Gingeras TR, *et al.* (2005) *Cell* 120:169–181.
24. Cao R, Wang L, Wang H, Xia L, Erdjument-Bromage H, Tempst P, Jones RS, Zhang Y (2002) *Science* 298:1039–1043.
25. Feng Q, Wang H, Ng HH, Erdjument-Bromage H, Tempst P, Struhl K, Zhang Y (2002) *Curr Biol* 12:1052–1058.
26. Ng HH, Feng Q, Wang H, Erdjument-Bromage H, Tempst P, Zhang Y, Struhl K (2002) *Genes Dev* 16:1518–1527.
27. Kouskouti A, Talianidis I (2005) *EMBO J* 24:347–357.
28. Maison C, Almouzni G (2004) *Nat Rev Mol Cell Biol* 5:296–304.
29. Melcher M, Schmid M, Aagaard L, Selenko P, Laible G, Jenuwein T (2000) *Mol Cell Biol* 20:3728–3741.
30. Nakayama J, Rice JC, Strahl BD, Allis CD, Grewal SI (2001) *Science* 292:110–113.
31. Vakoc CR, Mandat SA, Olenchock BA, Blobel GA (2005) *Mol Cell* 19:381–391.
32. Frasor J, Danes JM, Komm B, Chang KC, Lyttle CR, Katzenellenbogen BS (2003) *Endocrinology* 144:4562–4574.
33. Coser KR, Chesnes J, Hur J, Ray S, Isselbacher KJ, Shioda T (2003) *Proc Natl Acad Sci USA* 100:13994–13999.
34. Rae JM, Johnson MD, Scheys JO, Cordero KE, Larios JM, Lippman ME (2005) *Breast Cancer Res Treat* 92:141–149.
35. Zhu P, Baek SH, Bourk EM, Ohgi KA, Garcia-Bassets I, Sanjo H, Akira S, Kotol PF, Glass CK, Rosenfeld MG, Rose DW (2006) *Cell* 124:615–629.
36. Saville B, Wormke M, Wang F, Nguyen T, Enmark E, Kuiper G, Gustafsson JA, Safe S (2000) *J Biol Chem* 275:5379–5387.
37. Stein B, Yang MX (1995) *Mol Cell Biol* 15:4971–4979.
38. DeNardo DG, Kim HT, Hilsenbeck S, Cuba V, Tsimelzon A, Brown PH (2005) *Mol Endocrinol* 19:362–378.
39. Cheng AS, Jin VX, Fan M, Smith LT, Liyanarachchi S, Yan PS, Leu YW, Chan MW, Plass C, Nephew KP, *et al.* (2006) *Mol Cell* 21:393–404.
40. Miller LD, Smeds J, George J, Vega VB, Vergara L, Ploner A, Pawitan Y, Hall P, Klaar S, Liu ET, Bergh J (2005) *Proc Natl Acad Sci USA* 102:13550–13555.
41. Shang Y, Hu X, DiRenzo J, Lazar MA, Brown M (2000) *Cell* 103:843–852.
42. Yeakley JM, Fan JB, Doucet D, Luo L, Wickham E, Ye Z, Chee MS, Fu XD (2002) *Nat Biotechnol* 20:353–358.
43. Li Z, Van Calcar S, Qu C, Cavenee WK, Zhang MQ, Ren B (2003) *Proc Natl Acad Sci USA* 100:8164–8169.
44. Ren B, Dynlacht BD (2004) *Methods Enzymol* 376:304–315.
45. Tusher VG, Tibshirani R, Chu G (2001) *Proc Natl Acad Sci USA* 98:5116–5121.
46. Gibbons FD, Proft M, Struhl K, Roth FP (2005) *Genome Biol* 6:R96.